# SOLVING SYSTEMS OF LINEAR DIFFERENTIAL EQUATIONS BY WALSH POLYNOMIALS APPROACH

**Rodolfo Toledo** (Budapest, Hungary)

Communicated by László Szili

(Received January 29, 2021; accepted April 26, 2021)

**Abstract.** György Gát and the author of this paper established a procedure to solve initial value problems of linear differential equations of first-order with not necessary constant coefficients. This procedure approximates the exact solution by Walsh polynomials. This paper extends this method for systems of linear differential equations of first-order with discontinuous righthand sides.

## 1. Introduction

Orthonormal systems formed by Walsh functions have a wide range of applications in the world of digital technology. Since the 1970s several researchers studied intensively the application of Walsh functions in communication and signal processing (see, among others, [1, 13, 14, 16, 17]). In [7] Corrington developed a method to solve $n$th order linear differential equations using previously prepared huge tables of the Walsh–Fourier coefficients of weighted indefinite integrals of Walsh functions.

In 1975 C. F. Chen and C. H. Hsiao improved the method of Corrington to solve systems of linear differential equations of first-order. The new approach

*Key words and phrases*: Fourier analysis, Walsh–Paley system, numerical solution of linear differential equations, triangular functions, uniform convergence, differential equations with discontinuous righthand sides.
*2010 Mathematics Subject Classification*: 42C10, 65L05.

is much simpler and more suitable for digital computation, but it was only designed for solving linear differential equations with constant coefficients having initial conditions at $x = 0$. In 1975 Chen and Hsiao wrote many papers in which they show the performance of their procedure in different applications (see [6, 4, 5, 3]). On the basis of this method it was also possible to develop a technique for solving first-order partial differential equations by double Walsh series approximation (see [19]).

The basic idea of the method of Chen and Hsiao is to avoid differentiation considering the equivalent vector integral equation, because Walsh functions are not differentiable. They discretized the integral equations substituting all functions, even the integral functions, by the partial sums of Walsh series of them. Every component of the exact solution is also substituted by an unknown Walsh polynomial. The aim is to find the coefficients of these polynomials which are obtained after solving a linear system. The proposed numerical solution is formed by these Walsh polynomials which are defined in the interval $[0, 1[$.

In [12] György Gát and the author of this paper established a procedure to solve initial value problems of linear differential equations of first-order with not necessary constant coefficients. We used the same basic idea, but the proposed numerical solution was more complex and its analysis required a more solid mathematical background. We also propose an iterative algorithm to speed up the computations without having to solve very large linear systems. In this paper we extend these results for systems of linear differential equations.

As is known, the system of differential equations $y' = f(x, y)$ with a continuous right-hand side is equivalent to the vector integral equation

$$y(x) = y(x_0) + \int_{x_0}^{x} f(t, y(t))\, dt,$$

where $y(x)$ and $f(x, y)$ are vector-valued functions. If $f(x, y)$ is discontinuous in $x$ and continuous in $y$, then the functions satisfying the integral equation above can be called solutions of the equation $y' = f(x, y)$ (see [8]). Differential equations with discontinuous righthand sides have a large number of application. In addition, if this system of differential equations is linear having the initial condition $y(\xi) = \eta$, where $\xi$ belongs to an interval $I$, and its coefficients and free terms are integrable (but not necessarily continuous) on each closed subinterval of $I$, then there is an unique solution on $I$ which is obviously continuous (but not necessarily differentiable). Integration is to be understood in the sense of Lebesgue.

Our aim is to establish a procedure in order to approximate all members of this unique solution by Walsh polynomials. Since Walsh polynomials are defined on the interval $[0, 1[$, we suppose that $I = [0, 1]$. The approximation at the point $x = 1$ will be given by the continuity of the solution at $x = 1$

and the uniform convergence of the approximation on the interval $[0,1[$ which will be proved later in this paper. If we only need to find an approximation on a subinterval of $[0,1]$ then we can assume that all functions are zero on the complementer of this subinterval.

### 1.1. Motivation and main results

Chen and Hsiao did not deal with the extensive analysis of the proposed numerical solution. For instance they did not determine if the linear system is solvable or not, the convergence of the numerical solution or the estimation of errors. Moreover, to obtain a numeric solution with great accuracy it is necessary to solve a linear system with a very large number of equations. The construction of Walsh polynomials from their coefficients also requires a considerable amount of time.

In [10, 11] we started with the analysis of these issues considering the simplest case

$$y' + ay = q(x), \qquad y(0) = \eta,$$

where $a, \eta \in \mathbf{R}$ and the constant term $q$ is an integrable and continuous real function on $[0,1[$. We approximated the exact solution by a Walsh polynomial

$$\overline{y}_n(x) = \sum_{k=0}^{2^n-1} c_k w_k(x),$$

where $w_k$ is the $k$th Walsh function ordered in Paley's sense. For every $n$ the method gives us an unique numerical solution, except for $2^{n+1} = -a$. It was also proved that $\overline{y}_n$ converges uniformly to the exact solution $y$ if $n$ tends to infinity. In addition we proposed a faster iterative method to obtain directly values of the numerical solution without solving linear systems or constructing Walsh polynomials from their coefficients.

In [12] we established a similar method for differential equations with non-constant coefficient. We designed a new procedure to approach by Walsh polynomials the solution of the initial value problem

$$y' + p(x)y = q(x), \qquad y(0) = \eta,$$

where $\eta \in \mathbf{R}$ and $p, q$ are integrable and continuous real functions on $[0,1[$. The results were compatible with those for differential equations with constant coefficient. Namely, the proposed numerical solution $\overline{y}_n$ just cannot be constructed for finitely many values of $n$, and it converges uniformly to the exact solution of the initial value problem on the interval $[0,1[$, if $n$ tends to infinity. Moreover, a similar iterative method was given to speed up the computations. In [15]

Károly Nagy adapted this procedure for problems with the initial condition $y(\xi) = \eta$, where $0 < \xi < 1$.

The goal of the present work is to extend this method for linear systems of differential equations of first-order as general as possible. We note that the method solves the integral equation which is equivalent to the differential equation in case of all functions in the equation are continuous. The continuity ensures the existence of a differentiable solution, but otherwise the integral equation has an unique continuous solution on the interval $[0, 1]$ if all functions in the equation are integrable on this interval. Therefore, leaving aside the continuity, we consider the vector integral equation

$$y(x) = \eta + \int_{\xi}^{x} (q(t) - P(t)y(t))\, dt \qquad (x \in [0, 1]),$$

where $\xi \in [0, 1[$ and $P(x)$ and $q(x)$ are respectively a matrix and a vector with integrable entries on $[0, 1]$. Note that the initial condition is not necessarily $\xi = 0$ as it appears so far in our papers.

After discretizing the vector integral equation we find a vector $\overline{y}_n(x)$ formed by Walsh polynomials which satisfies it. To this end, we propose two different methods to construct these polynomials. One is to follow the method of Chen and Hsiao constructing the polynomials from their coefficients. It leads us to solve a very large linear system. The other is an iterative method like we proposed in [12] for one equation. In this paper we prove that an unique $\overline{y}_n$ always exists, except for finitely many values of $n$. Moreover, $\overline{y}_n$ converges uniformly to the exact solution on the interval $[0, 1]$, if $n$ tends to infinity. Finally we propose another representation of the numerical solution to obtain more accuracy in the approximation.

### 1.2. Structure of the paper

In Section 2.1 we introduce the notation used throughout this paper. We also formulate two theorems that are fundamental for the properly functioning of the proposed method. Further notations and a brief introduction to Walsh functions appear in Section 2.2. To learn more about Walsh functions see [18].

In Section 3 we propose two methods to construct the numerical solution. First, in 3.1 we describe how the coefficients of the Walsh polynomials in the solution may be obtained solving a large linear system. In 3.2 we introduce an iterative method which gives us directly the values of the numerical solution.

The implementation of this method in Matlab appears in Section 4 besides three solved examples. Here we modify the method to increase the accuracy of the approximation, comparing it with other classical numerical methods. Section 5 contains the proof of Theorems 2.1 and 2.2.

## 2. Preliminaries

### 2.1. Notation and results

Consider a general initial value problem with respect to a linear system of differential equations of first order on the interval $[0,1]$, that is

$$\begin{aligned}
y_1' + p_{1,1}(x)y_1 + p_{1,2}(x)y_2 + \cdots + p_{1,m}(x)y_m &= q_1(x), \qquad y_1(\xi) = \eta_1,\\
y_2' + p_{2,1}(x)y_1 + p_{2,2}(x)y_2 + \cdots + p_{2,m}(x)y_m &= q_2(x), \qquad y_2(\xi) = \eta_2,\\
&\cdots\\
y_m' + p_{m,1}(x)y_1 + p_{m,2}(x)y_2 + \cdots + p_{m,m}(x)y_m &= q_m(x), \quad y_m(\xi) = \eta_m,
\end{aligned}$$

where coefficients $p_{k,l}$ and free terms $q_k$ are defined on the interval $[0,1]$ for all $k,l = 1,2,\ldots,m$, and $\xi \in [0,1[$, $\eta_k \in \mathbf{R}$. A shorthand notation of this problem is

$$y' + P(x)y = q(x), \qquad y(\xi) = \eta, \tag{2.1}$$

where $P(x)$ and $q(x)$ are the matrix and the vector with entries $p_{k,l}(x)$ and $q_k(x)$ respectively, $\eta$ is the vector formed by the numbers $\eta_1$, $\eta_1$, …, $\eta_m$, and $y(x)$ denotes the vector formed by the solutions $y_1(x)$, $y_2(x)$,…, $y_m(x)$.

We suppose that the coefficients and free terms of the equations are integrable on the interval $[0,1]$, but not necessarily continuous. Therefore, the solution of (2.1) means the unique solution of the vector integral equation

$$y(x) = \eta + \int_{\xi}^{x} (q(t) - P(t)y(t))\, dt \qquad (x \in [0,1]) \tag{2.2}$$

which is continuous on $[0,1]$. We approximate every member of this solution by a Walsh polynomial on the interval $[0,1[$, more precisely, for every member $y_k$ we construct a polynomial

$$y_k^{(n)}(x) = \sum_{i=0}^{2^n-1} c_{k,i} w_i(x) \qquad (k = 1,2,\ldots,m), \tag{2.3}$$

where $w_i$ is the $i$th Walsh function ordered in Paley's sense.

In this regard we discretize the vector integral equation (2.2) as follows. Let $n$ be a positive integer. First consider the matrix

$$\mathcal{S}_{2^n} p(x) = \begin{pmatrix} S_{2^n}p_{1,1}(x) & S_{2^n}p_{1,2}(x) & \ldots & S_{2^n}p_{1,m}(x)\\ S_{2^n}p_{2,1}(x) & S_{2^n}p_{2,2}(x) & \ldots & S_{2^n}p_{2,m}(x)\\ \vdots & \vdots & & \vdots\\ S_{2^n}p_{m,1}(x) & S_{2^n}p_{m,2}(x) & \ldots & S_{2^n}p_{m,m}(x)\end{pmatrix},$$

and vectors

$$\overline{y}_n(x) = \begin{pmatrix} y_1^{(n)}(x) \\ y_2^{(n)}(x) \\ \vdots \\ y_m^{(n)}(x) \end{pmatrix}, \qquad S_{2^n}q(x) = \begin{pmatrix} S_{2^n}q_1(x) \\ S_{2^n}q_2(x) \\ \vdots \\ S_{2^n}q_m(x) \end{pmatrix},$$

where $S_{2^n}f(x)$ denotes the $2^n$-th partial sums of Walsh–Fourier series of the integrable function $f$ at the point $x \in [0,1[$. Note, that every function above is constant on the intervals

$$\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right[ \qquad (i = 0, 1, \ldots, 2^n - 1),$$

the so-called dyadic intervals of length $2^{-n}$. Then substitute $y(x)$, $P(x)$ and $q(x)$ by $\overline{y}_n(x)$, $\mathcal{S}_{2^n}p(x)$ and $S_{2^n}q(x)$ respectively in the integral equation. But it is not all, we must also discretize the integral function, otherwise the right-hand side is not constant on the dyadic intervals of length $2^{-n}$. For this purpose, we also take the $2^n$-th partial sums of Walsh–Fourier series of the integral function, using the expression

$$S_{2^n}\left(\int_\xi^{\cdot} f(t)\,dt\right)(x)$$

to denote the vector formed by the $2^n$-th partial sums of Walsh–Fourier series of the integral functions

$$\int_\xi^x f_k(t)\,dt \qquad (k = 1, 2, \ldots, m)$$

at the point $x$, where $f_k(x)$ are the components of the vector function $f(x)$.

In summary, we find a vector $\overline{y}_n(x)$ formed by Walsh polynomials which satisfies the discretized integral equation

$$(2.4) \quad \overline{y}_n(x) = \eta + S_{2^n}\left(\int_\xi^{\cdot} \left(S_{2^n}q(t) - \mathcal{S}_{2^n}p(t)\overline{y}_n(t)\right)\,dt\right)(x) \qquad (x \in [0,1[).$$

Note that (2.4) is equivalent to the equations

$$(2.5) \qquad y_k^{(n)}(x) = \eta_k + S_{2^n}\left(\int_\xi^{\cdot} \left(S_{2^n}q_k(t) - \sum_{l=1}^{m} S_{2^n}p_{k,l}(t)y_l^{(n)}(t)\right)dt\right)(x)$$

for all $k = 1, 2, \ldots, m$. We call $\overline{y}_n$ a numerical solution of the vector integral equation (2.2). In this paper we prove the following results.

**Theorem 2.1.** *Let $p_{k,l}$ and $q_k$ be integrable functions on the interval $[0,1]$ for all $k,l = 1,2,\dots,m$. Then there exists an unique vector $\overline{y}_n$ formed by Walsh polynomials which satisfies the discretized integral equation (2.4), except for finitely many values of $n$.*

In other words, an unique numerical solution $\overline{y}_n$ of the vector integral equation (2.2) exists with the exception of finitely many values of $n$.

**Theorem 2.2.** *Let $p_{k,l}$ and $q_k$ be integrable functions on the interval $[0,1]$ for all $k,l = 1,2,\dots,m$. Then the numerical solutions $\overline{y}_n$ converge uniformly to the exact solution of the problem (2.2) on the interval $[0,1[$.*

Note that the $\overline{y}_n$ can be extended at the point $x = 1$ with the limit from the left of 1 of the Walsh polynomials. Then the continuity of the exact solution and Theorem 2.2 ensure the uniform approximation on the interval $[0,1]$.

In later sections we are using the infinity norm of matrices. We recall that it is the maximum absolute row sum of the matrix. In other words, if $A$ is an $n \times m$ matrix with entries $a_{k,l}$ then denote

$$\|A\| := \max_{1\le k\le n} \left\{ \sum_{l=1}^{m} |a_{k,l}| \right\}.$$

Consequently, the norm of a vector as a matrix with only one column is

$$\|b\| := \max_{1\le k\le n} \{|b_k|\},$$

where $b_k$ are the components of the vector $b$. Aside from the common properties of a norm it is not hard to see that

$$\|AB\| \le \|A\|\|B\|,$$

particularly when $B$ is a matrix with only one column being, therefore, a vector. Moreover, if the matrix function $F(x)$ is integrable on the interval $I$, then the function $\|F(x)\|$ is also integrable on the interval $I$ and

$$\left\| \int_I F(x)\,dx \right\| \le \int_I \|F(x)\|\,dx.$$

## 2.2. The Walsh–Paley system

Walsh functions are the finite product of Rademacher functions. One of the ways used to introduce the Rademacher system is to consider the dyadic expansion $(x_0, x_1, \dots)$ of a real number $x \in [0,1[$ given by the sum

$$x = \sum_{k=0}^{\infty} \frac{x_k}{2^{k+1}},$$

where $x_k = 0$ or $x_k = 1$ for all $k \in \mathbf{N}$. This expansion is not unique if $x$ is a dyadic rational, that is, if $x$ is a number of the form $\frac{i}{2^k}$, where $i, k \in \mathbf{N}$ and $0 \leq i < 2^k$. When this situation occurs we choose the expansion terminating in zeros. Then, the Rademacher system is defined by

$$r_k(x) := (-1)^{x_k} \qquad (x \in [0,1[,\ k \in \mathbf{N}).$$

The Walsh–Paley system sort the Walsh functions as follows. Every $n \in \mathbf{N}$ can be uniquely expressed as

$$n = \sum_{k=0}^{\infty} n_k 2^k,$$

where $n_k = 0$ or $n_k = 1$ for all $k \in \mathbf{N}$. This allows us to say that the sequence $(n_0, n_1, \dots)$ is the dyadic expansion of $n$. Then, the Walsh–Paley system is defined by

$$w_n(x) := \prod_{k=0}^{\infty} r_k^{n_k}(x) \qquad (x \in [0,1[,\ n \in \mathbf{N}).$$

$w_n$ is called the $n$th Walsh–Paley function or in other words, the $n$th Walsh function ordered in Paley's sense. A Walsh polynomial is the finite linear combination of Walsh functions. The Walsh–Paley system is orthonormal, i.e.

$$\int_0^1 w_n(x) w_m(x)\, dx = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$$

Among other things, the orthonormality of the Walsh–Paley system ensures the fact that two Walsh polynomial are equal at every point if and only if they have the same coefficients.

For an integrable function $f$, i.e.

$$\int_0^1 |f(x)|\, dx < \infty,$$

we define the Fourier coefficients and partial sums of Walsh–Fourier series by

$$\widehat{f}_k := \int_0^1 f(x) w_k(x)\, dx \quad (k \in \mathbf{N}),$$

$$S_n f(x) := \sum_{k=0}^{n-1} \widehat{f}_k w_k(x) \quad (n \in \mathbf{N},\ x \in [0,1[).$$

The $2^n$-th partial sums of Walsh–Fourier series play a prominent role in applications, since their values can be written as an integral mean, namely

$$S_{2^n} f(x) = 2^n \int\limits_{I_n(x)} f(y)\, dy,$$

where the sets

$$I_k(i) := \left[\frac{i}{2^k}, \frac{i+1}{2^k}\right[ \qquad (i = 0, 1, \dots, 2^k - 1,\ k \in \mathbf{N})$$

are called dyadic intervals and $I_n(x)$ denotes the dyadic interval which contains the value of $x$. Moreover, $S_{2^n} f$ converges uniformly to $f$ for every continuous function $f$ defined on the interval $[0, 1]$, and $S_{2^n} f$ converges to $f$ in $L^1$-norm for every integrable function $f$ (see [18] p. 142).

Moduli of continuity helps us to estimate the approximations above. For this, we define first the dyadic sum of two numbers $x, y \in [0, 1[$ with expansion $(x_0, x_1, \dots)$ and $(y_0, y_1, \dots)$ respectively by

$$x \dotplus y := \sum_{k=0}^{\infty} |x_k - y_k| 2^{-(k+1)}.$$

The dyadic modulus of continuity of a continuous function $f$ defined on $[0, 1]$ is

$$\omega_n f := \sup\{|f(x \dotplus h) - f(x)| \colon x \in [0, 1[, 0 \le h < 2^{-n}\}$$

which tends to zero if $n$ tends to infinity. It is not hard to prove (see [18]), that

$$|S_{2^n} f(x) - f(x)| \le \omega_n f \qquad (x \in [0, 1[).$$

On the other hand, the dyadic $L^1$-modulus of continuity of every integrable function $f$ is defined by

$$\omega_n^{(1)} f := \sup \left\{ \int\limits_0^1 |f(x \dotplus h) - f(x)|\, dx \colon x \in [0, 1[, 0 \le h < 2^{-n} \right\}$$

which also tends to zero if $n$ tends to infinity. In this case (see [18])

$$\int\limits_0^1 |S_{2^n} f(x) - f(x)|\, dx \le \omega_n^{(1)} f.$$

In case of a matrix or vector of functions we can defined moduli of continuity as follows. Let $P(x)$ be a matrix with continuous entries $p_{k,l}(x)$ on $[0, 1]$ and $\Omega_n P$ be the matrix with entries $\omega_n p_{k,l}$. Then, the dyadic modulus of continuity of $P(x)$ is defined by

$$\omega_n P := \|\Omega_n P\|.$$

Similarly, for a matrix $P(x)$ with integrable entries $p_{k,l}(x)$ on $[0,1]$ let $\Omega_n^{(1)}P$ be the matrix with entries $\omega_n^{(1)}p_{k,l}$, and define

$$\omega_n^{(1)}P := \|\Omega_n^{(1)}P\|.$$

It is not hard to see that the corresponding estimations

$$\|S_{2^n}P(x) - P(x)\| \le \omega_n P \qquad \text{and} \qquad \left\| \int_0^1 |S_{2^n}P(x) - P(x)|\,dx \right\| \le \omega_n^{(1)}P$$

hold. It is important to clarify that in this paper the absolute value of a matrix means the matrix formed by the absolute value if its entries. All the above is valid for vector functions if we consider them as matrices with only one column.

Continuing our matrix notation we consider the integral function of the Walsh–Paley functions, the so-called triangular functions. We denote them by

$$J_i(x) := \int_0^x w_i(t)\,dt \qquad (i \in \mathbf{N},\ 0 \le x < 1).$$

It is obvious that $J_0(x) = x$. Moreover,

$$J_i(x) = w_i(x)(x - x^{(n)}),$$

where $2^n \le i < 2^{n+1}$ and $x^{(n)}$ is the number of the form $\frac{j}{2^n}$ for $j = 0, 1, \ldots, 2^n$ which is nearer to $x$ (see [9]). Denote by $\widehat{J}_{i,j}$ the $j$th Walsh–Fourier coefficient of the triangular function $J_i$. Let $\widehat{J}$ be the matrix whose entries are $\widehat{J}_{i,j}$, where $i, j = 0, 1, \ldots, 2^n - 1$. The entries of $\widehat{J}$ often take the value 0 and it is not hard to construct it by iteration (see [2] and [10]). For instance, for $n = 3$ we have

$$\widehat{J} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & -\frac{1}{8} & 0 & -\frac{1}{16} & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & -\frac{1}{8} & 0 & -\frac{1}{16} & 0 & 0 \\ \frac{1}{8} & 0 & 0 & 0 & 0 & 0 & -\frac{1}{16} & 0 \\ 0 & \frac{1}{8} & 0 & 0 & 0 & 0 & 0 & -\frac{1}{16} \\ \frac{1}{16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{16} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{16} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{16} & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.6}$$

Similarly let $\widehat{J}_\xi$ be the matrix of the Walsh–Fourier coefficient of the integral functions

$$\int_\xi^x w_i(t)\,dt \qquad (i \in \mathbf{N}, 0 \le x < 1).$$

To compute the values of the matrix $\widehat{J}_\xi$ note that

$$\int\limits_{\xi}^{x} w_i(t)\,dt = J_i(x) - J_i(\xi) = J_i(x) - w_i(\xi)(\xi - \xi^{(n)}).$$

Therefore, the matrix $\widehat{J}_\xi$ only differs from $\widehat{J}$ in its first column in which the value of the $i$th row is subtracted by $w_i(\xi)(\xi - \xi^{(n)})$.

## 3. Construction of the numerical solution

We propose two methods to construct the numerical solution $\overline{y}_n$. They are different as far as the calculation is concerned, but the generated numerical solution is the same. The first one is the extension of the method of Chen and Hsiao for equations with variable coefficients. In case of one equation the iterative method was introduced in [11] for constant coefficients and in [12] for variable coefficients.

### 3.1. Construction of the Walsh polynomials from their coefficients

Our aim is to obtain the coefficients $c_{k,i}$ of the Walsh polynomials

$$y_k^{(n)}(x) = \sum_{i=0}^{2^n-1} c_{k,i} w_i(x) \qquad (k = 1, 2, \dots, m)$$

which satisfy the discretized equations (2.5), that is

$$y_k^{(n)}(x) = \eta_k + S_{2^n}\left( \int\limits_{\xi}^{\cdot} S_{2^n} q_k(t) - \sum_{l=1}^{m} S_{2^n} p_{k,l}(t) y_l^{(n)}(t)\,dt \right)(x)$$

for all $k = 1, 2, \dots, m$, where $x \in [0, 1[$. Then, we construct them by the linear combination of Walsh functions.

This requires the introduction of the following matrix notations. First, define the vectors

$$\mathbf{c_k} := (c_{k,0}, c_{k,1}, \dots, c_{k,2^n-1})^\top \quad \text{and} \quad \boldsymbol{w}(x) := (w_0(x), w_1(x), \dots, w_{2^n-1}(x))^\top.$$

Then $y_k^{(n)}(x) = \boldsymbol{w}(x)^\top \mathbf{c_k}$. Let $\widehat{\mathbf{p_{k,l}}}$ and $\widehat{\mathbf{q_k}}$ be the vectors formed by the Walsh–Fourier coefficients of the functions $p_{k,l}$ and $q_k$ respectively. Therefore, the $i$th component of vectors $\widehat{\mathbf{p_{k,l}}}$ and $\widehat{\mathbf{q_k}}$ are

$$\widehat{\mathbf{p_{k,l}}}_i := \int\limits_0^1 p_{k,l}(t) w_i(t)\,dt \qquad \text{and} \qquad \widehat{\mathbf{q_k}}_i := \int\limits_0^1 q_k(t) w_i(t)\,dt.$$

The concept of partial sums of Walsh–Fourier series means that

$$S_{2^n} p_{k,l}(x) = \boldsymbol{w}(x)^\top \widehat{\mathbf{p_{k,l}}} \qquad \text{and} \qquad S_{2^n} q_k(x) = \boldsymbol{w}(x)^\top \widehat{\mathbf{q_k}}.$$

With the vectors $\widehat{\mathbf{p_{k,l}}}$ we construct the matrix

$$\widehat{P_{k,l}} := \left(\widehat{\mathbf{p_{k,l}}}_{i\oplus j}\right)_{i,j=0}^{2^n-1},$$

where $i \oplus j$ is the dyadic sum of the integers $i$ and $j$. Note that $\widehat{P_{k,l}}$ is the dyadically circulant matrix generated by the coefficients of $S_{2^n} p_{k,l}(x)$ (see [12]). In addition, denote $\mathbf{e_0} := (1, 0, \ldots, 0)^\top$ the first unit vector of size $2^n$. Since $w_0(x) = 1$ for all $x \in [0, 1[$ we have $\eta_k = \boldsymbol{w}(x)^\top \eta_k \mathbf{e_0}$.

With the matrix notations above (see also the definition of $\widehat{J_\xi}$ in the previous section) the discretized integral equations (2.5) can be written as follows

$$\begin{aligned}
\boldsymbol{w}(x)^\top \mathbf{c_k} &= \eta_k + S_{2^n}\left(\int_\xi^{\cdot} \left(\boldsymbol{w}(t)^\top \widehat{\mathbf{q_k}} - \sum_{l=1}^m \boldsymbol{w}(t)^\top \widehat{\mathbf{p_{k,l}}} \boldsymbol{w}(t)^\top \mathbf{c_l}\right) dt\right)(x) = \\
&= \eta_k + S_{2^n}\left(\int_\xi^{\cdot} \left(\boldsymbol{w}(t)^\top \widehat{\mathbf{q_k}} - \boldsymbol{w}(t)^\top \sum_{l=1}^m \widehat{P_{k,l}} \mathbf{c_l}\right) dt\right)(x) = \\
&= \boldsymbol{w}(x)^\top \eta_k \mathbf{e_0} + S_{2^n}\left(\int_\xi^{\cdot} \boldsymbol{w}(t)^\top \, dt\right)(x) \cdot \left(\widehat{\mathbf{q_k}} - \sum_{l=1}^m \widehat{P_{k,l}} \mathbf{c_l}\right) = \qquad (3.1) \\
&= \boldsymbol{w}(x)^\top \eta_k \mathbf{e_0} + \boldsymbol{w}(x)^\top \widehat{J_\xi}^\top \left(\widehat{\mathbf{q_k}} - \sum_{l=1}^m \widehat{P_{k,l}} \mathbf{c_l}\right) = \\
&= \boldsymbol{w}(x)^\top \left(\eta_k \mathbf{e_0} + \widehat{J_\xi}^\top \left(\widehat{\mathbf{q_k}} - \sum_{l=1}^m \widehat{P_{k,l}} \mathbf{c_l}\right)\right).
\end{aligned}$$

at every point of $[0, 1[$. In the equation above we used the following result

$$\begin{aligned}
\sum_{l=1}^m \boldsymbol{w}(t)^\top \widehat{\mathbf{p_{k,l}}} \boldsymbol{w}(t)^\top \mathbf{c_l} &= \sum_{l=1}^m \left(\sum_{r=0}^{2^n-1} \widehat{\mathbf{p_{k,l}}}_r w_r(t) \sum_{j=0}^{2^n-1} c_{l,j} w_j(t)\right) = \\
&= \sum_{l=1}^m \sum_{r,j=0}^{2^n-1} \widehat{\mathbf{p_{k,l}}}_r c_{l,j} w_r(t) w_j(t) = \sum_{l=1}^m \sum_{r,j=0}^{2^n-1} \widehat{\mathbf{p_{k,l}}}_r c_{l,j} w_{r\oplus j}(t) = \\
&= \sum_{l=1}^m \sum_{i,j=0}^{2^n-1} \widehat{\mathbf{p_{k,l}}}_{i\oplus j} c_{l,j} w_i(t) = \boldsymbol{w}(t)^\top \sum_{l=1}^m \widehat{P_{k,l}} \mathbf{c_l}
\end{aligned}$$

by the fact that the dyadic sum is a group operation and if $i = r \oplus j$ then $i = 0, 1, \ldots, 2^n - 1$ and $r = i \oplus j$.

Both sides of the equation (3.1) are Walsh polynomials, hence their coefficients are the same, that is

$$\mathbf{c_k} = \eta_k \mathbf{e_0} + \widehat{J}_\xi^\top \left( \widehat{\mathbf{q_k}} - \sum_{l=1}^{m} \widehat{P_{k,l}} \mathbf{c_l} \right).$$

This can be written as follows

$$\mathbf{c_k} + \sum_{l=1}^{m} \widehat{J}_\xi^\top \widehat{P_{k,l}} \mathbf{c_l} = \eta_k \mathbf{e_0} + \widehat{J}_\xi^\top \widehat{\mathbf{q_k}}. \tag{3.2}$$

Taking into consideration that (3.2) is valid for all $k = 1, 2, \ldots, m$, we have obtained a linear system of $m2^n$ equations involving the variables

$$c_{1,0}, c_{1,1}, \ldots, c_{1,2^n-1}, c_{2,0}, c_{2,1}, \ldots, c_{2,2^n-1}, \ldots, c_{m,0}, c_{m,1}, \ldots, c_{m,2^n-1}.$$

In order to write this linear system in only one matrix equation denote by $\mathbf{c}$ the vector formed by the variables above, by $\widehat{\mathcal{J}}_\xi$ the Kronecker product of the identity matrix of size $m$ and $\widehat{J}_\xi$, and by $\boldsymbol{\eta}$ the Kronecker product of $\eta$ and $\mathbf{e_0}$. We also denote the following block matrices

$$\widehat{P} := \begin{pmatrix} \widehat{P_{1,1}} & \widehat{P_{1,2}} & \ldots & \widehat{P_{1,m}} \\ \widehat{P_{2,1}} & \widehat{P_{2,2}} & \ldots & \widehat{P_{2,m}} \\ \vdots & \vdots & \ldots & \vdots \\ \widehat{P_{m,1}} & \widehat{P_{m,2}} & \ldots & \widehat{P_{m,m}} \end{pmatrix} \quad \text{and} \quad \widehat{\mathbf{q}} := \begin{pmatrix} \widehat{q_1} \\ \widehat{q_2} \\ \vdots \\ \widehat{q_m} \end{pmatrix}.$$

Finally we denote by $\mathcal{I}$ the identity matrix of size $m2^n$. Then, the linear system which gives us the coefficients of the numerical solution is

$$(\mathcal{I} + \widehat{\mathcal{J}}_\xi^\top P)\mathbf{c} = \boldsymbol{\eta} + \widehat{\mathcal{J}}_\xi^\top \widehat{\mathbf{q}}. \tag{3.3}$$

### 3.2. The iterative method

The disadvantage of the construction of the Walsh polynomials by its coefficients is the requirement of solving a liner system with a very large number of equations. In addition, we must construct the Walsh polynomials with the computed coefficients. The amount of time required for these computations would really be large if we try to obtain a high accuracy. In this section we propose a faster method for directly getting the values of the numerical solution without needing to solve the linear systems and generate Walsh polynomials.

The method is based on the fact that the numerical solution $\overline{y}_n$ is constant on the dyadic intervals of length $2^{-n}$, i.e. for all $i = 0, 1, \dots, 2^n - 1$ we have

$$\overline{y}_n(x) = \overline{y}_n\Big(\frac{i}{2^n}\Big) \qquad \Big(\frac{i}{2^n} \le x < \frac{i+1}{2^n}\Big).$$

The point is to calculate iteratively the value of all $\overline{y}_n(\frac{i}{2^n})$ starting from the value of $\overline{y}_n(\xi)$. First, we generalize Lemma 1 of [12].

**Lemma 3.1.** *Suppose $f\colon [0,1[\to \mathbf{R}^m$ is constant on the dyadic intervals of length $2^{-n}$ and $i^*$ is the integer such that $\frac{i^*}{2^n} \le \xi < \frac{i^*+1}{2^n}$. Denote*

$$d_n(\xi) := i^* - 2^n\xi + \frac{1}{2}.$$

*For $x \in [0,1[$ let $i$ be the integer such that $\frac{i}{2^n} \le x < \frac{i+1}{2^n}$. Then*

$$S_{2^n}\Bigg(\int\limits_{\xi}^{\cdot} f(t)\,dt\Bigg)(x) = \begin{cases} \frac{d_n(\xi)}{2^n} f(\frac{i^*}{2^n}), & i = i^*, \\ \frac{d_n(\xi)+\frac{1}{2}}{2^n} f(\frac{i^*}{2^n}) + \frac{1}{2^n}\sum\limits_{j=i^*+1}^{i-1} f(\frac{j}{2^n}) + \frac{1}{2^{n+1}} f(\frac{i}{2^n}), & i > i^*, \\ \frac{d_n(\xi)-\frac{1}{2}}{2^n} f(\frac{i^*}{2^n}) - \frac{1}{2^n}\sum\limits_{j=i+1}^{i^*-1} f(\frac{j}{2^n}) - \frac{1}{2^{n+1}} f(\frac{i}{2^n}), & i < i^*. \end{cases}$$

**Proof.** The operator $S_{2^n}f$ is linear, therefore

$$S_{2^n}\Bigg(\int\limits_{\xi}^{\cdot} f(t)\,dt\Bigg)(x) = S_{2^n}\Bigg(\int\limits_{0}^{\cdot} f(t)\,dt - \int\limits_{0}^{\xi} f(t)\,dt\Bigg)(x) =$$

$$= S_{2^n}\Bigg(\int\limits_{0}^{\cdot} f(t)\,dt\Bigg)(x) - \int\limits_{0}^{\xi} f(t)\,dt.$$

The first term was studied in [11] where it was proved that

$$S_{2^n}\Bigg(\int\limits_{0}^{\cdot} f(t)\,dt\Bigg)(x) = \frac{1}{2^n}\sum_{j=0}^{i-1} f\Big(\frac{j}{2^n}\Big) + \frac{1}{2^{n+1}} f\Big(\frac{i}{2^n}\Big).$$

Since $f$ is constant on the dyadic intervals of length $2^{-n}$, the second term is

$$\int\limits_{0}^{\xi} f(t)\,dt = \sum_{j=0}^{i^*-1} \int\limits_{\frac{j}{2^n}}^{\frac{j+1}{2^n}} f(t)\,dt + \int\limits_{\frac{i^*}{2^n}}^{\xi} f(t)\,dt = \frac{1}{2^n}\sum_{j=0}^{i^*-1} f\Big(\frac{j}{2^n}\Big) + f\Big(\frac{i^*}{2^n}\Big)\Big(\xi - \frac{i^*}{2^n}\Big) =$$

$$= \frac{1}{2^n}\sum_{j=0}^{i^*-1} f\Big(\frac{j}{2^n}\Big) - \frac{d_n(\xi)-\frac{1}{2}}{2^n} f\Big(\frac{i^*}{2^n}\Big).$$

Subtracting the results above we obtain directly the statement of the lemma. ■

In the lemma above we introduce the number $d_n(\xi)$ for which is easy to see that $-\frac{1}{2} < d_n(\xi) \leq \frac{1}{2}$ and $\xi = \frac{1}{2}$ if and only if $\xi$ can be written of the form $\xi = \frac{i}{2^n}$.

Now back to the discretized integral equation (2.4), that is

$$\overline{y}_n(x) = \eta + S_{2^n}\left(\int_{\xi}^{\cdot} (S_{2^n}q(t) - \mathcal{S}_{2^n}p(t)\overline{y}_n(t))\, dt\right)(x) \qquad (x \in [0,1[).$$

Note that functions appeared in the integral function are constant on the dyadic intervals of length $2^{-n}$, therefore we may directly apply the lemma on these equations to obtain the following cases:

- if $i = i^*$ then

$$\overline{y}_n\Big(\frac{i^*}{2^n}\Big) = \eta + \frac{d_n(\xi)}{2^n}\left(S_{2^n}q\Big(\frac{i^*}{2^n}\Big) - \mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\overline{y}_n\Big(\frac{i^*}{2^n}\Big)\right),$$

- if $i > i^*$ then

$$\begin{aligned}\overline{y}_n\Big(\frac{i}{2^n}\Big) = \eta &+ \frac{d_n(\xi)+\frac{1}{2}}{2^n}\left(S_{2^n}q\Big(\frac{i^*}{2^n}\Big) - \mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\overline{y}_n\Big(\frac{i^*}{2^n}\Big)\right) + \\ &+ \frac{1}{2^n}\sum_{j=i^*+1}^{i-1}\left(S_{2^n}q\Big(\frac{j}{2^n}\Big) - \mathcal{S}_{2^n}p\Big(\frac{j}{2^n}\Big)\overline{y}_n\Big(\frac{j}{2^n}\Big)\right) + \\ &+ \frac{1}{2^{n+1}}\left(S_{2^n}q\Big(\frac{i}{2^n}\Big) - \mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\overline{y}_n\Big(\frac{i}{2^n}\Big)\right),\end{aligned}$$

- if $i < i^*$ then

$$\begin{aligned}\overline{y}_n\Big(\frac{i}{2^n}\Big) = \eta &+ \frac{d_n(\xi)-\frac{1}{2}}{2^n}\left(S_{2^n}q\Big(\frac{i^*}{2^n}\Big) - \mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\overline{y}_n\Big(\frac{i^*}{2^n}\Big)\right) - \\ &- \frac{1}{2^n}\sum_{j=i+1}^{i^*-1}\left(S_{2^n}q\Big(\frac{j}{2^n}\Big) - \mathcal{S}_{2^n}p\Big(\frac{j}{2^n}\Big)\overline{y}_n\Big(\frac{j}{2^n}\Big)\right) - \\ &- \frac{1}{2^{n+1}}\left(S_{2^n}q\Big(\frac{i}{2^n}\Big) - \mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\overline{y}_n\Big(\frac{i}{2^n}\Big)\right).\end{aligned}$$

Consequently,

- if $i = i^*$ then

$$\left(\mathcal{I} + \frac{d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\right)\overline{y}_n\Big(\frac{i^*}{2^n}\Big) = \eta + \frac{d_n(\xi)}{2^n}S_{2^n}q\Big(\frac{i^*}{2^n}\Big),$$

- if $i > i^*$ then

$$\left(\mathcal{I} + \frac{1}{2^{n+1}}\mathcal{S}_{2^n} p\left(\frac{i}{2^n}\right)\right)\overline{y}_n\left(\frac{i}{2^n}\right) = \eta +$$
$$+ \frac{d_n(\xi)+\frac{1}{2}}{2^n}\left(S_{2^n} q\left(\frac{i^*}{2^n}\right) - \mathcal{S}_{2^n} p\left(\frac{i^*}{2^n}\right)\overline{y}_n\left(\frac{i^*}{2^n}\right)\right) +$$
$$+ \frac{1}{2^n}\sum_{j=i^*+1}^{i-1}\left(S_{2^n} q\left(\frac{j}{2^n}\right) - \mathcal{S}_{2^n} p\left(\frac{j}{2^n}\right)\overline{y}_n\left(\frac{j}{2^n}\right)\right) + \frac{1}{2^{n+1}}S_{2^n} q\left(\frac{i}{2^n}\right),$$

- if $i < i^*$ then

$$\left(\mathcal{I} - \frac{1}{2^{n+1}}\mathcal{S}_{2^n} p\left(\frac{i}{2^n}\right)\right)\overline{y}_n\left(\frac{i}{2^n}\right) = \eta +$$
$$+ \frac{d_n(\xi)-\frac{1}{2}}{2^n}\left(S_{2^n} q\left(\frac{i^*}{2^n}\right) - \mathcal{S}_{2^n} p\left(\frac{i^*}{2^n}\right)\overline{y}_n\left(\frac{i^*}{2^n}\right)\right) -$$
$$- \frac{1}{2^n}\sum_{j=i+1}^{i^*-1}\left(S_{2^n} q\left(\frac{j}{2^n}\right) - \mathcal{S}_{2^n} p\left(\frac{j}{2^n}\right)\overline{y}_n\left(\frac{j}{2^n}\right)\right) - \frac{1}{2^{n+1}}S_{2^n} q\left(\frac{i}{2^n}\right),$$

where $\mathcal{I}$ is the identity matrix of size $m$.

These formulae give us the following bidirectional iterative method to obtain the values of the numerical solution $\overline{y}_n$.

- The method starts with the calculation of the values of the vector $\overline{y}_n(\frac{i^*}{2^n})$ solving the linear system related to the case $i = i^*$.
- With the values of $\overline{y}_n(\frac{i^*}{2^n})$ we solve the linear system related to the case $i > i^*$ for $i = i^* + 1$ to obtain the vector $\overline{y}_n(\frac{i^*+1}{2^n})$.
- With the values of $\overline{y}_n(\frac{i^*}{2^n})$ and $\overline{y}_n(\frac{i^*+1}{2^n})$ we solve the linear system related to the case $i > i^*$ for $i = i^* + 2$ to obtain the vector $\overline{y}_n(\frac{i^*+2}{2^n})$.
- We follow the process above for $i = i^* + 3, i^* + 4, \ldots, 2^n - 1$ to obtain the vectors $\overline{y}_n(\frac{i^*+3}{2^n})$, $\overline{y}_n(\frac{i^*+4}{2^n})$,...,$\overline{y}_n(\frac{2^n-1}{2^n})$.
- if $i^* > 1$ then with the values of $\overline{y}_n(\frac{i^*-1}{2^n})$ we solve the linear system related to the case $i < i^*$ for $i = i^* - 1$ to obtain the vector $\overline{y}_n(\frac{i^*-1}{2^n})$.
- With the values of $\overline{y}_n(\frac{i^*}{2^n})$ and $\overline{y}_n(\frac{i^*-1}{2^n})$ we solve the linear system related to the case $i < i^*$ for $i = i^* - 2$ to obtain the vector $\overline{y}_n(\frac{i^*-2}{2^n})$.
- We follow the process above for $i = i^* - 3, i^* - 4, \ldots, 0$ to obtain the vectors $\overline{y}_n(\frac{i^*-3}{2^n})$, $\overline{y}_n(\frac{i^*-4}{2^n})$,...,$\overline{y}_n(0)$.

In every step we must solve a linear system of of $m$ variables and $m$ equations.

## 4. Implementation of the iterative method in Matlab

### 4.1. Code implementation

In Matlab it is not difficult to implement the iterative algorithm using a vectorized code. The following Matlab function expects the value of $n$ and the parameters of the differential equation in matrix form, and it returns a matrix having in its rows the value of the members of the numerical solution on the dyadic intervals of length $2^{-n}$. It also returns the left endpoints of these intervals. Note that the passed matrix function `PNumF` and vector function `qNumF` should be defined as a function handle for which the use of the built-in function `matlabFunction` may result very useful if we are working with symbolic expressions.

```
function [x,yNum]=odeLinWalsh(n,PNumF,qNumF,xi,eta,atol,rtol)
% Solving linear initial value problems using Walsh functions
% by the iterative method

% Input variables
% n: it sets the dimension of the solution
% PNumF: matrix of coefficients (function handle)
% qNumF: vector of free terms (function handle)
% xi: initial point
% eta: exact values at the initial point
% atol: absolute tolerance used only for numerical integration
% rtol: relative tolerance used only for numerical integration

% Auxiliary variables
dim = 2^n;                    % dimension of the solution
m = length(eta);              % number of equations
iStart = floor(xi*dim)+1;     % index of the initial point
dev = iStart-dim*xi;          % dev = d_n(xi)-1/2
Id = eye(m);                  % identity matrix of size m

% Output variables
yNum = zeros(m,dim);          % values of the numerical solution
x = 0:1/dim:1-1/dim;          % left endpoints of diadic intervals

% Computation of the integral means
FSp = cell(1,dim);
FSq = cell(1,dim);
for i = 1:dim
    FSp{i} = integral(PNumF,(i-1)/dim,i/dim,'ArrayValued',true,...
                            'AbsTol',atol,'RelTol',rtol);
    FSq{i} = integral(qNumF,(i-1)/dim,i/dim,'ArrayValued',true,...
                            'AbsTol',atol,'RelTol',rtol);
end
```

```

% Value at the starting index
A = Id+(dev-1/2)*FSp{iStart};
b = eta+(dev-1/2)*FSq{iStart};
yNum(:,iStart) = linsolve(A,b);

% Forward computations
szum = eta+dev*(FSq{iStart}-FSp{iStart}*yNum(:,iStart));
for k = iStart+1:dim
    A = Id+1/2*FSp{k};
    b = szum+1/2*FSq{k};
    yNum(:,k) = linsolve(A,b);
    szum = szum+FSq{k}-FSp{k}*yNum(:,k);
end

% Backward computations
szum = eta+(dev-1)*(FSq{iStart}-FSp{iStart}*yNum(:,iStart));
for k = iStart-1:-1:1
    A = Id-1/2*FSp{k};
    b = szum-1/2*FSq{k};
    yNum(:,k) = linsolve(A,b);
    szum = szum-FSq{k}+FSp{k}*yNum(:,k);
end
end
```

### 4.2. The modified solution

Let us see an example: consider the problem

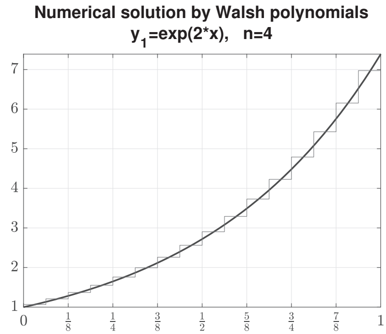
$$y_1' - 2y_1 + \frac{1}{x+2} y_2 = 2 - x, \qquad\qquad y_1(0) = 1,$$

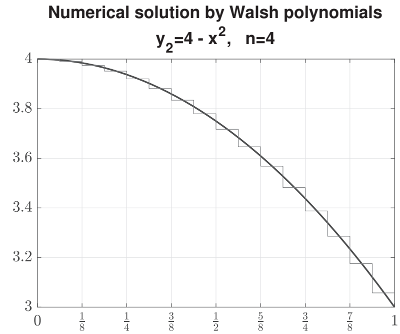
$$y_2' + xy_1 - \frac{2}{x-2} y_2 = xe^{2x} + 4, \qquad\qquad y_1(0) = 4.$$

on the interval $[0,1]$. It is not hard to verify that the exact solution of the problem is

$$y_1(x) = e^{2x}, \qquad y_2(x) = 4 - x^2.$$

After executing our algorithm for $n = 4$ we obtain the result illustrated in the following graphs.




To provide more details about the approximation we divide the interval $[0, 1]$ in subintervals of length $\frac{1}{8}$, and we compute the maximal difference between the exact and the numerical solution on these intervals for different values of $n$. The results appear in the following two tables. Observe that the maximal difference is reduced almost by half when the value of $n$ increases by one.

Maximal difference with respect to $y_1$

| $n$ | $[0, \frac{1}{8}[$ | $[\frac{1}{8}, \frac{1}{4}[$ | $[\frac{1}{4}, \frac{3}{8}[$ | $[\frac{3}{8}, \frac{1}{2}[$ | $[\frac{1}{2}, \frac{5}{8}[$ | $[\frac{5}{8}, \frac{3}{4}[$ | $[\frac{3}{4}, \frac{7}{8}[$ | $[\frac{7}{8}, 1[$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.01e+00 | 7.24e-01 | 3.60e-01 | 7.10e-01 | 3.35e+00 | 2.58e+00 | 1.59e+00 | 1.32e+00 |
| 2 | 3.34e-01 | 3.15e-01 | 5.76e-01 | 4.94e-01 | 9.91e-01 | 7.72e-01 | 1.70e+00 | 1.20e+00 |
| 3 | 1.43e-01 | 1.86e-01 | 2.41e-01 | 3.13e-01 | 4.06e-01 | 5.26e-01 | 6.83e-01 | 8.86e-01 |
| 4 | 7.58e-02 | 9.78e-02 | 1.26e-01 | 1.63e-01 | 2.10e-01 | 2.72e-01 | 3.51e-01 | 4.53e-01 |
| 5 | 3.90e-02 | 5.02e-02 | 6.46e-02 | 8.32e-02 | 1.07e-01 | 1.38e-01 | 1.78e-01 | 2.29e-01 |
| 6 | 1.98e-02 | 2.54e-02 | 3.27e-02 | 4.20e-02 | 5.41e-02 | 6.95e-02 | 8.94e-02 | 1.15e-01 |
| 7 | 9.96e-03 | 1.28e-02 | 1.64e-02 | 2.11e-02 | 2.71e-02 | 3.49e-02 | 4.48e-02 | 5.76e-02 |
| 8 | 5.00e-03 | 6.42e-03 | 8.25e-03 | 1.06e-02 | 1.36e-02 | 1.75e-02 | 2.24e-02 | 2.88e-02 |
| 9 | 2.50e-03 | 3.21e-03 | 4.13e-03 | 5.30e-03 | 6.81e-03 | 8.75e-03 | 1.12e-02 | 1.44e-02 |
| 10 | 1.25e-03 | 1.61e-03 | 2.07e-03 | 2.65e-03 | 3.41e-03 | 4.37e-03 | 5.62e-03 | 7.21e-03 |

Maximal difference with respect to $y_2$

| $n$ | $[0, \frac{1}{8}[$ | $[\frac{1}{8}, \frac{1}{4}[$ | $[\frac{1}{4}, \frac{3}{8}[$ | $[\frac{3}{8}, \frac{1}{2}[$ | $[\frac{1}{2}, \frac{5}{8}[$ | $[\frac{5}{8}, \frac{3}{4}[$ | $[\frac{3}{4}, \frac{7}{8}[$ | $[\frac{7}{8}, 1[$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.17e-01 | 1.02e-01 | 5.50e-02 | 1.33e-01 | 5.12e-01 | 3.72e-01 | 2.00e-01 | 2.38e-01 |
| 2 | 2.91e-02 | 3.34e-02 | 9.01e-02 | 9.74e-02 | 1.61e-01 | 1.52e-01 | 2.56e-01 | 1.82e-01 |
| 3 | 8.11e-03 | 2.43e-02 | 4.02e-02 | 5.56e-02 | 7.04e-02 | 8.77e-02 | 1.07e-01 | 1.29e-01 |
| 4 | 5.97e-03 | 1.39e-02 | 2.18e-02 | 2.95e-02 | 3.71e-02 | 4.55e-02 | 5.43e-02 | 6.37e-02 |
| 5 | 3.45e-03 | 7.38e-03 | 1.13e-02 | 1.52e-02 | 1.91e-02 | 2.31e-02 | 2.73e-02 | 3.16e-02 |
| 6 | 1.84e-03 | 3.80e-03 | 5.75e-03 | 7.70e-03 | 9.65e-03 | 1.16e-02 | 1.37e-02 | 1.57e-02 |
| 7 | 9.48e-04 | 1.93e-03 | 2.90e-03 | 3.88e-03 | 4.85e-03 | 5.84e-03 | 6.83e-03 | 7.84e-03 |
| 8 | 4.81e-04 | 9.69e-04 | 1.46e-03 | 1.95e-03 | 2.43e-03 | 2.92e-03 | 3.42e-03 | 3.91e-03 |
| 9 | 2.42e-04 | 4.86e-04 | 7.30e-04 | 9.74e-04 | 1.22e-03 | 1.46e-03 | 1.71e-03 | 1.95e-03 |
| 10 | 1.22e-04 | 2.44e-04 | 3.66e-04 | 4.88e-04 | 6.10e-04 | 7.32e-04 | 8.54e-04 | 9.77e-04 |

Most of the numerical method for solving differential equations determine the value of the solution at some points and use linear interpolation to calculate the solution at other points. In this regard we modify our proposed numerical solution as follows. Our original solution consists of a step function which is constant on dyadic intervals. From now on, we only keep the value of the solution at the middle of these intervals, and then we use linear interpolation to connect these points. But the first half of the first interval and the second half of the last interval fall outside. We continue straight the line of the interpolation next to these half intervals to extend the interpolation to the whole interval $[0,1]$, unless the initial value is at $\xi = 0$. In this case we ad the point $(\xi, \eta)$ to the set of interpolation points.

In the way described above we get a broken line, consisting of little straight line segments. We call it the modified solution of our method. This modification increase significantly the accuracy of the approximation. The maximal difference between the exact and the modified solution reveals this fact. Observe in the table below that now the maximal difference is reduced by almost a quarter when the value of $n$ increases by one.

Maximal difference with respect to $y_1$

| $n$ | $[0,\frac{1}{8}[$ | $[\frac{1}{8},\frac{1}{4}[$ | $[\frac{1}{4},\frac{3}{8}[$ | $[\frac{3}{8},\frac{1}{2}[$ | $[\frac{1}{2},\frac{5}{8}[$ | $[\frac{5}{8},\frac{3}{4}[$ | $[\frac{3}{4},\frac{7}{8}[$ | $[\frac{7}{8},1[$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 2.20e-01 | 3.60e-01 | 9.07e-01 | 1.32e+00 | 1.56e+00 | 1.61e+00 | 1.59e+00 | 1.33e+00 |
| 2 | 5.00e-02 | 1.31e-01 | 1.36e-01 | 2.49e-01 | 2.61e-01 | 4.66e-01 | 4.90e-01 | 4.32e-01 |
| 3 | 2.23e-02 | 3.09e-02 | 4.27e-02 | 5.86e-02 | 8.01e-02 | 1.09e-01 | 1.48e-01 | 1.50e-01 |
| 4 | 5.50e-03 | 7.63e-03 | 1.05e-02 | 1.44e-02 | 1.97e-02 | 2.69e-02 | 3.65e-02 | 4.26e-02 |
| 5 | 1.37e-03 | 1.90e-03 | 2.62e-03 | 3.60e-03 | 4.91e-03 | 6.69e-03 | 9.08e-03 | 1.14e-02 |
| 6 | 3.42e-04 | 4.75e-04 | 6.55e-04 | 8.98e-04 | 1.23e-03 | 1.67e-03 | 2.27e-03 | 2.96e-03 |
| 7 | 8.55e-05 | 1.19e-04 | 1.64e-04 | 2.24e-04 | 3.07e-04 | 4.18e-04 | 5.67e-04 | 7.54e-04 |
| 8 | 2.14e-05 | 2.97e-05 | 4.09e-05 | 5.61e-05 | 7.67e-05 | 1.04e-04 | 1.42e-04 | 1.90e-04 |
| 9 | 5.35e-06 | 7.42e-06 | 1.02e-05 | 1.40e-05 | 1.92e-05 | 2.61e-05 | 3.54e-05 | 4.78e-05 |
| 10 | 1.34e-06 | 1.86e-06 | 2.56e-06 | 3.51e-06 | 4.79e-06 | 6.52e-06 | 8.86e-06 | 1.20e-05 |

Maximal difference with respect to $y_2$

| $n$ | $[0,\frac{1}{8}[$ | $[\frac{1}{8},\frac{1}{4}[$ | $[\frac{1}{4},\frac{3}{8}[$ | $[\frac{3}{8},\frac{1}{2}[$ | $[\frac{1}{2},\frac{5}{8}[$ | $[\frac{5}{8},\frac{3}{4}[$ | $[\frac{3}{4},\frac{7}{8}[$ | $[\frac{7}{8},1[$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 4.31e-02 | 5.52e-02 | 1.38e-01 | 1.90e-01 | 2.10e-01 | 2.11e-01 | 2.00e-01 | 1.58e-01 |
| 2 | 1.36e-02 | 2.84e-02 | 2.83e-02 | 3.18e-02 | 3.20e-02 | 5.20e-02 | 5.62e-02 | 5.25e-02 |
| 3 | 7.26e-03 | 7.25e-03 | 6.86e-03 | 7.29e-03 | 8.66e-03 | 1.14e-02 | 1.61e-02 | 1.66e-02 |
| 4 | 1.88e-03 | 1.81e-03 | 1.70e-03 | 1.78e-03 | 2.10e-03 | 2.75e-03 | 3.87e-03 | 4.73e-03 |
| 5 | 4.78e-04 | 4.51e-04 | 4.25e-04 | 4.44e-04 | 5.22e-04 | 6.81e-04 | 9.60e-04 | 1.28e-03 |
| 6 | 1.21e-04 | 1.13e-04 | 1.06e-04 | 1.11e-04 | 1.30e-04 | 1.70e-04 | 2.39e-04 | 3.35e-04 |
| 7 | 3.04e-05 | 2.82e-05 | 2.65e-05 | 2.77e-05 | 3.26e-05 | 4.25e-05 | 5.98e-05 | 8.57e-05 |
| 8 | 7.61e-06 | 7.05e-06 | 6.63e-06 | 6.92e-06 | 8.14e-06 | 1.06e-05 | 1.49e-05 | 2.17e-05 |
| 9 | 1.90e-06 | 1.76e-06 | 1.66e-06 | 1.73e-06 | 2.04e-06 | 2.65e-06 | 3.74e-06 | 5.46e-06 |
| 10 | 4.77e-07 | 4.40e-07 | 4.14e-07 | 4.33e-07 | 5.09e-07 | 6.63e-07 | 9.34e-07 | 1.37e-06 |

Let us just do a quick comparison with other classical numerical methods, namely the Euler's, the midpoint and the classical Runge–Kutta method. We compare the maximal differences between the exact and these numerical solutions on the whole interval $[0,1]$. The results appear in the table below, where $n$ means that the step size is $h = 2^{-n}$. We note that the accuracy of the modified solution is similar to those of midpoint method.

Maximal differences on the interval $[0,1]$

| $n$ | Modified | | Euler | | Midpoint | | Runge–Kutta | |
|---|---|---|---|---|---|---|---|---|
| | $y_1$ | $y_2$ | $y_1$ | $y_2$ | $y_1$ | $y_2$ | $y_1$ | $y_2$ |
| 1 | 1.61e+00 | 2.11e-01 | 3.44e+00 | 5.13e-01 | 1.19e+00 | 2.00e-01 | 5.38e-01 | 6.30e-02 |
| 2 | 4.90e-01 | 5.62e-02 | 2.38e+00 | 3.94e-01 | 4.36e-01 | 8.81e-02 | 1.77e-01 | 1.57e-02 |
| 3 | 1.50e-01 | 1.66e-02 | 1.47e+00 | 2.63e-01 | 1.33e-01 | 2.79e-02 | 5.07e-02 | 3.91e-03 |
| 4 | 4.26e-02 | 4.73e-03 | 8.33e-01 | 1.55e-01 | 3.69e-02 | 7.79e-03 | 1.35e-02 | 9.77e-04 |
| 5 | 1.14e-02 | 1.28e-03 | 4.46e-01 | 8.44e-02 | 9.67e-03 | 2.05e-03 | 3.50e-03 | 2.44e-04 |
| 6 | 2.96e-03 | 3.35e-04 | 2.31e-01 | 4.42e-02 | 2.48e-03 | 5.26e-04 | 8.88e-04 | 6.10e-05 |
| 7 | 7.54e-04 | 8.57e-05 | 1.18e-01 | 2.26e-02 | 6.27e-04 | 1.33e-04 | 2.24e-04 | 1.53e-05 |
| 8 | 1.90e-04 | 2.17e-05 | 5.94e-02 | 1.14e-02 | 1.58e-04 | 3.35e-05 | 5.62e-05 | 3.81e-06 |
| 9 | 4.78e-05 | 5.46e-06 | 2.99e-02 | 5.75e-03 | 3.95e-05 | 8.40e-06 | 1.41e-05 | 9.54e-07 |
| 10 | 1.20e-05 | 1.37e-06 | 1.50e-02 | 2.88e-03 | 9.88e-06 | 2.10e-06 | 3.52e-06 | 2.38e-07 |

It is important to note that our algorithm runs rather fast in spite of computing $2^n$ number of integrals. For instance, the calculation for $n = 10$ were made in 0.751 seconds, while for $n = 4$ only 0.013 seconds were necessary.

## 4.3. Solving a stiff problem

Let us see another example: consider the problem

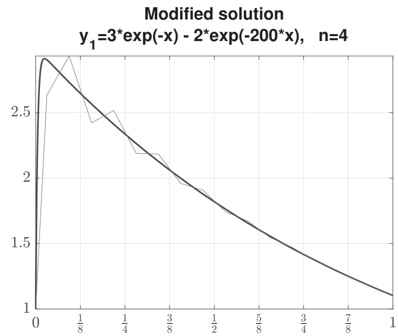
$$y_1' + \frac{806}{10}y_1 - \frac{1194}{10}y_2 = 0, \qquad\qquad y_1(0) = 1,$$

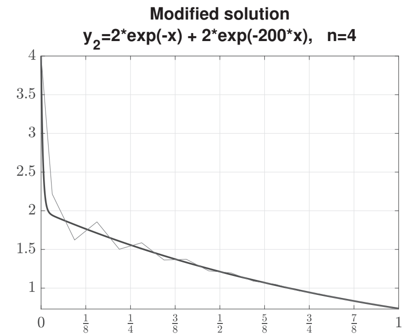
$$y_2' - \frac{707}{10}y_1 - \frac{1204}{10}y_2 = 0, \qquad\qquad y_1(0) = 4.$$

on the interval $[0,1]$. The exact solution of the problem is

$$y_1(x) = 3e^{-x} - 2e^{-200x}, \qquad y_2(x) = 2e^{-x} + 2e^{-200x}.$$

Both are functions with a very high slope in a neighborhood of the point $x = 0$. The following graphs show us the result of the approximation of the modified solution for $n = 4$.




As a consequence of this behavior, the maximal difference does not decay as rapid as we saw in the section above. This is seen clearly in the following tables.

Maximal difference with respect to $y_1$

| $n$ | $[0, \frac{1}{8}[$ | $[\frac{1}{8}, \frac{1}{4}[$ | $[\frac{1}{4}, \frac{3}{8}[$ | $[\frac{3}{8}, \frac{1}{2}[$ | $[\frac{1}{2}, \frac{5}{8}[$ | $[\frac{5}{8}, \frac{3}{4}[$ | $[\frac{3}{4}, \frac{7}{8}[$ | $[\frac{7}{8}, 1[$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.80e+00 | 9.67e-01 | 7.81e-02 | 9.96e-02 | 1.00e-01 | 9.27e-02 | 6.06e-02 | 6.75e-02 |
| 2 | 1.67e+00 | 5.77e-02 | 8.32e-02 | 8.32e-02 | 5.82e-02 | 5.82e-02 | 6.46e-02 | 9.53e-02 |
| 3 | 1.46e+00 | 1.30e-01 | 1.04e-01 | 9.43e-02 | 7.59e-02 | 6.81e-02 | 5.54e-02 | 9.43e-02 |
| 4 | 1.18e+00 | 1.44e-01 | 7.50e-02 | 3.91e-02 | 2.03e-02 | 1.05e-02 | 5.41e-03 | 3.26e-03 |
| 5 | 8.28e-01 | 3.39e-02 | 2.17e-03 | 5.20e-04 | 3.76e-04 | 3.11e-04 | 2.60e-04 | 3.68e-04 |
| 6 | 4.85e-01 | 1.62e-04 | 1.31e-04 | 1.10e-04 | 9.26e-05 | 7.76e-05 | 6.49e-05 | 9.09e-05 |
| 7 | 2.57e-01 | 3.87e-05 | 3.27e-05 | 2.75e-05 | 2.31e-05 | 1.94e-05 | 1.62e-05 | 2.26e-05 |
| 8 | 1.26e-01 | 9.68e-06 | 8.17e-06 | 6.88e-06 | 5.78e-06 | 4.85e-06 | 4.05e-06 | 5.63e-06 |
| 9 | 4.73e-02 | 2.42e-06 | 2.04e-06 | 1.72e-06 | 1.45e-06 | 1.21e-06 | 1.01e-06 | 1.40e-06 |
| 10 | 1.49e-02 | 6.05e-07 | 5.11e-07 | 4.30e-07 | 3.62e-07 | 3.03e-07 | 2.53e-07 | 3.50e-07 |

Maximal difference with respect to $y_2$

| $n$ | $[0, \frac{1}{8}[$ | $[\frac{1}{8}, \frac{1}{4}[$ | $[\frac{1}{4}, \frac{3}{8}[$ | $[\frac{3}{8}, \frac{1}{2}[$ | $[\frac{1}{2}, \frac{5}{8}[$ | $[\frac{5}{8}, \frac{3}{4}[$ | $[\frac{3}{4}, \frac{7}{8}[$ | $[\frac{7}{8}, 1[$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.81e+00 | 1.05e+00 | 8.67e-02 | 8.54e-02 | 6.77e-02 | 3.10e-02 | 9.06e-02 | 1.72e-01 |
| 2 | 1.68e+00 | 8.97e-02 | 6.29e-02 | 6.29e-02 | 7.05e-02 | 7.05e-02 | 5.78e-02 | 1.42e-01 |
| 3 | 1.47e+00 | 1.23e-01 | 1.10e-01 | 8.98e-02 | 7.94e-02 | 6.54e-02 | 5.74e-02 | 1.05e-01 |
| 4 | 1.18e+00 | 1.45e-01 | 7.65e-02 | 4.03e-02 | 2.12e-02 | 1.12e-02 | 5.95e-03 | 5.78e-03 |
| 5 | 8.28e-01 | 3.43e-02 | 2.56e-03 | 3.49e-04 | 2.42e-04 | 2.07e-04 | 1.73e-04 | 2.45e-04 |
| 6 | 4.85e-01 | 1.03e-04 | 8.72e-05 | 7.34e-05 | 6.17e-05 | 5.17e-05 | 4.32e-05 | 6.06e-05 |
| 7 | 2.57e-01 | 2.58e-05 | 2.18e-05 | 1.84e-05 | 1.54e-05 | 1.29e-05 | 1.08e-05 | 1.51e-05 |
| 8 | 1.26e-01 | 6.45e-06 | 5.45e-06 | 4.59e-06 | 3.86e-06 | 3.23e-06 | 2.70e-06 | 3.75e-06 |
| 9 | 4.73e-02 | 1.61e-06 | 1.36e-06 | 1.15e-06 | 9.64e-07 | 8.08e-07 | 6.76e-07 | 9.36e-07 |
| 10 | 1.49e-02 | 4.03e-07 | 3.40e-07 | 2.87e-07 | 2.41e-07 | 2.02e-07 | 1.69e-07 | 2.33e-07 |

Note that the results of the approximation are as expected outside the interval $\left[0, \frac{1}{8}\right]$. The significance of this observation is the fact that for this problem many of the classical numerical methods fail in a neighborhood of the point $x = 1$ using the step size $h = 2^{-4}$. For instance, with the classical Runge–Kutta method we obtain the very huge maximal difference of `2.39e+46`.

We can solve the problem in Matlab using the built-in function `ode23s` which is based on a modified Rosenbrock formula of order 2. It is a low order method to solve stiff differential equations. With a tolerance of 0.001 we obtain the following approximation based on 27 non-equidistant points.

Maximal difference using `ode23s` with tolerance 0.001

| $n$ | $\left[0, \frac{1}{8}\right[$ | $\left[\frac{1}{8}, \frac{1}{4}\right[$ | $\left[\frac{1}{4}, \frac{3}{8}\right[$ | $\left[\frac{3}{8}, \frac{1}{2}\right[$ | $\left[\frac{1}{2}, \frac{5}{8}\right[$ | $\left[\frac{5}{8}, \frac{3}{4}\right[$ | $\left[\frac{3}{4}, \frac{7}{8}\right[$ | $\left[\frac{7}{8}, 1\right[$ |
|---|---|---|---|---|---|---|---|---|
| $y_1$ | 8.12e-03 | 3.23e-03 | 2.47e-03 | 2.15e-03 | 1.88e-03 | 1.63e-03 | 1.23e-03 | 1.06e-03 |
| $y_2$ | 8.12e-03 | 2.16e-03 | 1.65e-03 | 1.44e-03 | 1.25e-03 | 1.09e-03 | 8.20e-04 | 7.10e-04 |

The results are comparable to those of the modified method for $n = 5$, except on the interval $\left[0, \frac{1}{8}\right]$.

### 4.4. An example with discontinuities

Finally we solve a problem with discontinuous and unbounded functions. Consider the problem

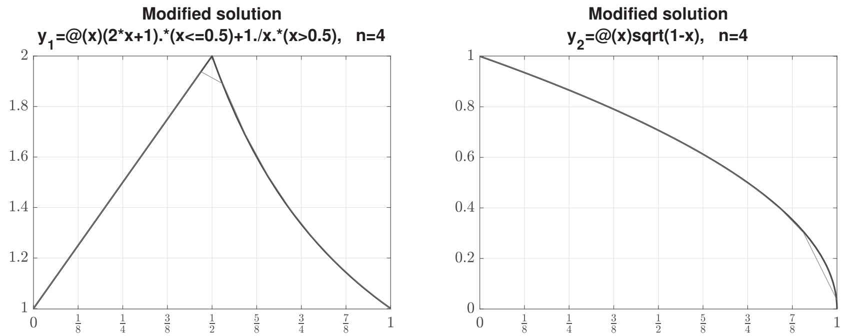
$$y_1' + xy_1 + \sqrt{1-x}\,y_2 = \begin{cases} 2x^2+3, & x \le \frac{1}{2}, \\ 2 - x - \frac{1}{x^2}, & x > \frac{1}{2}, \end{cases} \qquad y_1\left(\frac{3}{4}\right) = \frac{4}{3},$$

$$y_2' + \left(\begin{cases} \frac{1}{2x+1}, & x \le \frac{1}{2}, \\ x, & x > \frac{1}{2}, \end{cases}\right) y_1 - \frac{1}{\sqrt{1-x}} y_2 = -\frac{1}{2\sqrt{1-x}}, \qquad y_1\left(\frac{3}{4}\right) = \frac{1}{2},$$

on the interval $[0, 1]$. Note that all functions above are integrable on $[0, 1]$. The generalized concept of solution (see [8]) leads us to solve the related integral equation. The exact solution is

$$y_1(x) = \begin{cases} 2x+1, & x \le \frac{1}{2}, \\ \frac{1}{x}, & x > \frac{1}{2}, \end{cases} \qquad y_2(x) = \sqrt{1-x} \qquad (x \in [0, 1]).$$

Note that $y_1$ is continuous, but not differentiable at $x = \frac{1}{2}$. The following graphs illustrate the approximation of the modified solution for $n = 4$.




We can see that the accuracy of the approximation is worst in a neighborhood of the points $x = \frac{1}{2}$ and $x = 1$ with respect to other points. This is due to the discontinuity of the functions $p_{2,1}$ and $q_1$ at $x = \frac{1}{2}$ and also that the functions $p_{2,2}$ and $q_2$ are not bounded in a neighborhood of $x = 1$. This fact is shown in the following tables.

Maximal difference with respect to $y_1$

| $n$ | $[0, \frac{1}{8}[$ | $[\frac{1}{8}, \frac{1}{4}[$ | $[\frac{1}{4}, \frac{3}{8}[$ | $[\frac{3}{8}, \frac{1}{2}[$ | $[\frac{1}{2}, \frac{5}{8}[$ | $[\frac{5}{8}, \frac{3}{4}[$ | $[\frac{3}{4}, \frac{7}{8}[$ | $[\frac{7}{8}, 1[$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 2.94e-01 | 1.93e-01 | 4.36e-01 | 6.80e-01 | 6.80e-01 | 2.73e-01 | 1.97e-01 | 3.47e-01 |
| 2 | 8.93e-03 | 9.30e-03 | 9.68e-03 | 2.84e-01 | 2.84e-01 | 9.25e-02 | 8.71e-02 | 8.20e-02 |
| 3 | 2.16e-03 | 2.25e-03 | 2.36e-03 | 1.60e-01 | 1.60e-01 | 3.44e-02 | 1.92e-02 | 1.22e-02 |
| 4 | 5.33e-04 | 5.57e-04 | 5.86e-04 | 8.62e-02 | 8.63e-02 | 8.30e-03 | 4.67e-03 | 2.88e-03 |
| 5 | 1.33e-04 | 1.39e-04 | 1.46e-04 | 4.49e-02 | 4.49e-02 | 2.06e-03 | 1.16e-03 | 7.15e-04 |
| 6 | 3.32e-05 | 3.47e-05 | 3.65e-05 | 2.29e-02 | 2.29e-02 | 5.14e-04 | 2.90e-04 | 1.78e-04 |
| 7 | 8.30e-06 | 8.68e-06 | 9.13e-06 | 1.16e-02 | 1.16e-02 | 1.28e-04 | 7.23e-05 | 4.46e-05 |
| 8 | 2.08e-06 | 2.17e-06 | 2.28e-06 | 5.83e-03 | 5.83e-03 | 3.21e-05 | 1.81e-05 | 1.11e-05 |
| 9 | 5.19e-07 | 5.43e-07 | 5.70e-07 | 2.92e-03 | 2.92e-03 | 8.02e-06 | 4.52e-06 | 2.79e-06 |
| 10 | 1.30e-07 | 1.36e-07 | 1.43e-07 | 1.46e-03 | 1.46e-03 | 2.00e-06 | 1.13e-06 | 6.96e-07 |

Maximal difference with respect to $y_2$

| $n$ | $[0, \frac{1}{8}[$ | $[\frac{1}{8}, \frac{1}{4}[$ | $[\frac{1}{4}, \frac{3}{8}[$ | $[\frac{3}{8}, \frac{1}{2}[$ | $[\frac{1}{2}, \frac{5}{8}[$ | $[\frac{5}{8}, \frac{3}{4}[$ | $[\frac{3}{4}, \frac{7}{8}[$ | $[\frac{7}{8}, 1[$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 5.28e-02 | 2.52e-02 | 1.42e-02 | 2.28e-02 | 2.34e-02 | 2.02e-02 | 5.43e-02 | 3.15e-01 |
| 2 | 1.75e-02 | 8.21e-03 | 5.35e-03 | 5.35e-03 | 6.33e-03 | 7.38e-02 | 1.09e-01 | 1.09e-01 |
| 3 | 4.12e-03 | 1.80e-03 | 1.47e-03 | 1.15e-03 | 3.34e-03 | 8.49e-03 | 5.29e-02 | 7.50e-02 |
| 4 | 1.01e-03 | 4.71e-04 | 3.86e-04 | 3.09e-04 | 8.24e-04 | 1.99e-03 | 6.24e-03 | 5.23e-02 |
| 5 | 2.50e-04 | 1.21e-04 | 9.89e-05 | 7.96e-05 | 2.05e-04 | 4.91e-04 | 1.46e-03 | 3.67e-02 |
| 6 | 6.21e-05 | 3.05e-05 | 2.50e-05 | 2.02e-05 | 5.13e-05 | 1.22e-04 | 3.60e-04 | 2.59e-02 |
| 7 | 1.55e-05 | 7.68e-06 | 6.29e-06 | 5.09e-06 | 1.28e-05 | 3.05e-05 | 8.98e-05 | 1.83e-02 |
| 8 | 3.86e-06 | 1.93e-06 | 1.58e-06 | 1.28e-06 | 3.20e-06 | 7.63e-06 | 2.24e-05 | 1.29e-02 |
| 9 | 9.65e-07 | 4.83e-07 | 3.95e-07 | 3.20e-07 | 8.01e-07 | 1.91e-06 | 5.60e-06 | 9.15e-03 |
| 10 | 2.41e-07 | 1.21e-07 | 9.89e-08 | 8.00e-08 | 2.00e-07 | 4.77e-07 | 1.40e-06 | 6.47e-03 |

In case of $y_1$ the maximal difference is reduced almost by half on the interval $\left[\frac{3}{8}, \frac{5}{8}\right[$ when the value of $n$ increases by one. With respect to $y_2$ the maximal difference slowly decays on the interval $\left[\frac{7}{8}, 1\right[$. On the other intervals the maximal difference is reduced almost by a quarter as expected.

## 5. Proofs of the theorems

### 5.1. The existence of the numerical solution

In this section we prove Theorem 2.1, that is, we show that the numerical solution always exists, except for finitely many numbers of $n$. To this end, we going to study the solvability of the linear systems in the steps of the iterative method. It depends on whether the matrices

$$\begin{aligned} &\mathcal{I} + \frac{d_n(\xi)}{2^n}\mathcal{S}_{2^n} p\Big(\frac{i^*}{2^n}\Big) && (\text{for } i = i^*), \\ &\mathcal{I} + \frac{1}{2^{n+1}}\mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big) && (\text{for } i > i^*), \\ &\mathcal{I} - \frac{1}{2^{n+1}}\mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big) && (\text{for } i < i^*) \end{aligned} \tag{5.1}$$

are non-singular.

The entries of the matrix

$$\frac{1}{2^n}\mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big)$$

tend uniformly to zero if $n$ tends to infinity. Indeed, for every integrable function $f$ the integral function

$$F(x) = \int_0^x p(t)\,dt$$

is continuous on the closed interval $[0, 1]$, hence $F$ is also uniformly continuous. Therefore, for all $\varepsilon > 0$ exists a $\delta > 0$ such that for arbitrary $x_1, x_2 \in [0, 1]$ if $|x_1 - x_2| < \delta$ then $|F(x_1) - F(x_2)| < \varepsilon$. Thus, if $n_0$ is an index such that $\frac{1}{2^{n_0}} < \delta$ holds, then

$$\left|\frac{1}{2^n} S_{2^n} f\Big(\frac{i}{2^n}\Big)\right| = \left|\int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} f(x)\,dx\right| = \left|F\Big(\frac{i+1}{2^n}\Big) - F\Big(\frac{i}{2^n}\Big)\right| < \varepsilon$$

for all $n > n_0$ and $i = 0, 1, \ldots, 2^n - 1$. Since $\mathcal{S}_{2^n} p$ is composed of finitely many numbers of integrable functions, the sequence

$$L_n := \max_{\substack{0 \le i < 2^n \\ 1 \le k,l \le m}} \left\{\frac{1}{2^n}\left|S_{2^n} p_{k,l}\Big(\frac{i}{2^n}\Big)\right|\right\} \to 0 \qquad (\text{if } n \to \infty). \tag{5.2}$$

On the other hand, $|d_n(\xi)| \leq \frac{1}{2}$. This means that all matrices in (5.1) are "nearly identity" matrices. More precisely, there exists a positive integer $n_0$ such that if $n > n_0$, then for all matrices in (5.1) the absolute value of the diagonal entry in a row is larger than the sum of the absolute values of all the other entries in that row. Matrices having this property are called strictly diagonally dominant matrices. However, these matrices are non-singular. This result is known as the Levy–Desplanques theorem. Consequently, all matrices in (5.1) are non-singular, except for finite many values of $n$. This proves Theorem 2.1.

### 5.2. The uniform convergence of the numerical solution

The analysis of the approximation will be done according to the estimation

$$\|y(x) - \overline{y}_n(x)\| \leq \|S_{2^n} y(x) - y(x)\| + \|S_{2^n} y(x) - \overline{y}_n(x)\|. \tag{5.3}$$

First, recall that every member of the solution $y$ is a continuous function on the interval $[0,1]$. Hence the solution $y$ has finite modulus of continuity and

$$\|S_{2^n} y(x) - y(x)\| \leq \omega_n y \qquad (x \in [0,1[).$$

Therefore, the first addend of the right hand side of (5.3) tends uniformly to zero. For the estimation of the second addend we introduce the function

$$z_n(x) := \overline{y}_n(x) - S_{2^n} y(x) \qquad (x \in [0,1[).$$

By (2.4) and (2.2) we obtain

$$\begin{aligned} z_n(x) = \eta + S_{2^n}\left(\int_\xi^{\cdot} (S_{2^n} q(t) - \mathcal{S}_{2^n} p(t)\overline{y}_n(t))\ dt\right)(x) - \\ - S_{2^n}\left(\eta + \int_\xi^{\cdot} (q(t) - P(t)y(t))\ dt\right)(x) = \\ = S_{2^n}\left(\int_\xi^{\cdot} (S_{2^n} q(t) - q(t))\ dt\right)(x) - S_{2^n}\left(\int_\xi^{\cdot} (\mathcal{S}_{2^n} p(t) - P(t))y(t)\, dt\right)(x) + \\ + S_{2^n}\left(\int_\xi^{\cdot} \mathcal{S}_{2^n} p(t)(y(t) - S_{2^n} y(t))\, dt\right)(x) - S_{2^n}\left(\int_\xi^{\cdot} \mathcal{S}_{2^n} p(t) z_n(t)\, dt\right)(x). \end{aligned}$$

For simplicity we use the notation

$$\begin{aligned} m_n(x) := & S_{2^n}\left(\int_\xi^{\cdot} (S_{2^n} q(t) - q(t))\ dt\right)(x) - S_{2^n}\left(\int_\xi^{\cdot} (\mathcal{S}_{2^n} p(t) - P(t))y(t)\, dt\right)(x) + \\ & + S_{2^n}\left(\int_\xi^{\cdot} \mathcal{S}_{2^n} p(t)(y(t) - S_{2^n} y(t))\, dt\right)(x), \end{aligned} \tag{5.4}$$

therefore

$$z_n(x) = m_n(x) - S_{2^n}\left(\int_\xi^{\cdot} \mathcal{S}_{2^n} p(t) z_n(t)\, dt\right)(x). \tag{5.5}$$

Let $i^*$ be the positive integer such that $\frac{i^*}{2^n} \le \xi < \frac{i^*+1}{2^n}$ holds. For an $x \in [0,1[$ denoted by $i$ the positive integer such that $\frac{i}{2^n} \le x < \frac{i+1}{2^n}$ holds. Since $z_n$ is constant on these intervals, it is sufficient to consider only the values of $z_n(\frac{i}{2^n})$ for $i = 0, 1, \ldots, 2^n - 1$. Lemma 5.1 proves that

$$z_n\Big(\frac{i^*}{2^n}\Big) = \Big(\mathcal{I} + \frac{d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\Big)^{-1} m_n\Big(\frac{i^*}{2^n}\Big),$$

$$z_n\Big(\frac{i}{2^n}\Big) = \Big(\mathcal{I} + \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\Big)^{-1}\Big(m_n\Big(\frac{i}{2^n}\Big) - \sum_{j=i^*}^{i-1}\Big(\prod_{s=j+1}^{i-1}\big(\mathcal{I} - \rho_s^{(n)}\big)\Big)\rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big)\Big)$$

for all $i > i^*$, and

$$z_n\Big(\frac{i}{2^n}\Big) = \Big(\mathcal{I} - \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\Big)^{-1}\Big(m_n\Big(\frac{i}{2^n}\Big) - \sum_{j=i+1}^{i^*}\Big(\prod_{s=j+1}^{i-1}\big(\mathcal{I} - \rho_s^{(n)}\big)\Big)\rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big)\Big),$$

for all $i < i^*$, where

$$\rho_i^{(n)} = \begin{cases} \dfrac{d_n(\xi)+\frac{1}{2}}{2^n}\mathcal{S}_{2^n}p\Big(\dfrac{i^*}{2^n}\Big)\Big(\mathcal{I} + \dfrac{d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\dfrac{i^*}{2^n}\Big)\Big)^{-1}, & i = i^*, \\ \dfrac{1}{2^n}\mathcal{S}_{2^n}p\Big(\dfrac{i}{2^n}\Big)\Big(\mathcal{I} + \dfrac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\dfrac{i}{2^n}\Big)\Big)^{-1}, & i > i^*, \\ \dfrac{-1}{2^n}\mathcal{S}_{2^n}p\Big(\dfrac{i}{2^n}\Big)\Big(\mathcal{I} - \dfrac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\dfrac{i}{2^n}\Big)\Big)^{-1}, & i < i^*. \end{cases} \tag{5.6}$$

In Section 5.1 we proved the existence of the inverse matrices above with the exception of finitely many values of $n$.

For the estimation of the norm of $z_n$ we introduce the sequence

$$M_n := \omega_n^{(1)}q + \|y\|_\infty \omega_n^{(1)}P + 2mL_n\omega_n y,$$

where $\omega_n^{(1)}q$ and $\omega_n^{(1)}P$ denote the dyadic $L^1$ modulus of continuity of the vector of functions $q(x)$ and the matrix of functions $P(x)$ respectively, and $\omega_n y$ denotes the dyadic modulus of continuity of the exact solution $y(x)$ (see Section 2.2). Moreover, in the definition of $M_n$, $m$ is the number of equations in the initial value problem (2.1),

$$\|y\|_\infty := \max_{x\in[0,1]}\{\|y(x)\|\}$$

and

$$L_n = \max_{\substack{0\le i<2^n \\ 1\le k,l\le m}}\Big\{\frac{1}{2^n}\Big|S_{2^n}p_{k,l}\Big(\frac{i}{2^n}\Big)\Big|\Big\}$$

was already introduced in (5.2). Note that $M_n$ tends to zero if $n$ tends to infinity.

Lemma 5.2 ensures that $\|m_n(\frac{i}{2^n})\| \le M_n$ for all $i = 0, 1, \ldots, 2^n - 1$. By the basic properties of the norms and Lemma 5.3 we obtain directly from Lemma 5.1 that

$$\Big\|z_n\Big(\frac{i^*}{2^n}\Big)\Big\| \le \Big\|\Big(\mathcal{I} + \frac{d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\Big)^{-1}\Big\| M_n \le 2M_n.$$

Moreover, if $i > i^*$ then

$$\left\|z_n\Big(\frac{i}{2^n}\Big)\right\| \le \left\|\Big(\mathcal{I} + \frac{1}{2^{n+1}}\mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big)\Big)^{-1}\right\| \left(M_n + \sum_{j=i^*}^{i-1}\Big(\prod_{s=j+1}^{i-1}\Big(1+\|\rho_s^{(n)}\|\Big)\Big)\|\rho_j^{(n)}\|M_n\right) \le$$

$$\le 2M_n\left(1+\sum_{j=i^*}^{i-1}\|\rho_j^{(n)}\|\prod_{s=j+1}^{i-1}\Big(1+\|\rho_s^{(n)}\|\Big)\right) = 2M_n\prod_{j=i^*}^{i-1}\Big(1+\|\rho_j^{(n)}\|\Big).$$

In the calculations above we use the equality

$$1+\sum_{j=i^*}^{i-1}\|\rho_j^{(n)}\|\prod_{s=j+1}^{i-1}\Big(1+\|\rho_s^{(n)}\|\Big) = \prod_{j=i^*}^{i-1}\Big(1+\|\rho_j^{(n)}\|\Big)$$

which can be easily proved by using iteratively the transformation

$$1+\sum_{j=i^*}^{i-1}\|\rho_j^{(n)}\|\prod_{s=j+1}^{i-1}\Big(1+\|\rho_s^{(n)}\|\Big) = 1+\|\rho_{i-1}^{(n)}\|+\sum_{j=i^*}^{i-2}\|\rho_j^{(n)}\|\prod_{s=j+1}^{i-1}\Big(1+\|\rho_s^{(n)}\|\Big) =$$

$$= 1+\|\rho_{i-1}^{(n)}\|+(1+\|\rho_{i-1}^{(n)}\|)\sum_{j=i^*}^{i-2}\|\rho_j^{(n)}\|\prod_{s=j+1}^{i-2}\Big(1+\|\rho_s^{(n)}\|\Big) =$$

$$= (1+\|\rho_{i-1}^{(n)}\|)\left(1+\sum_{j=i^*}^{i-2}\|\rho_j^{(n)}\|\prod_{s=j+1}^{i-2}\Big(1+\|\rho_s^{(n)}\|\Big)\right).$$

In the same way we obtain for $i < i^*$ that

$$\left\|z_n\Big(\frac{i}{2^n}\Big)\right\| \le 2M_n\prod_{j=i+1}^{i^*}\Big(1+\|\rho_j^{(n)}\|\Big).$$

According to the definition of $\rho_i^{(n)}$ in (5.6) by Lemma 5.3 we have

$$\|\rho_{i^*}^{(n)}\| \le 2\left\|\frac{d_n(\xi)+\frac{1}{2}}{2^n}\mathcal{S}_{2^n} p\Big(\frac{i^*}{2^n}\Big)\right\| \le 2\left\|\frac{1}{2^n}\mathcal{S}_{2^n} p\Big(\frac{i^*}{2^n}\Big)\right\| \le 2\int_{\frac{i^*}{2^n}}^{\frac{i^*+1}{2^n}}\|P(t)\|\,dt,$$

since $0 < d_n(\xi)+\frac{1}{2} \le 1$. Similarly, for $i \ne i^*$ we have

$$\|\rho_i^{(n)}\| \le 2\left\|\frac{1}{2^n}\mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big)\right\| \le 2\left\|\frac{1}{2^n}\mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big)\right\| \le 2\int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\|P(t)\|\,dt.$$

Note that we used the upper estimation

$$\left\|\frac{1}{2^n}\mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big)\right\| = \left\|\int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} P(t)\,dt\right\| \le \int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\|P(t)\|\,dt,$$

since obviously $\|P(t)\|$ is integrable.

Therefore, by the inequality of arithmetic and geometric means we obtain

$$\|z_n(\frac{i}{2^n})\| \le 2M_n \prod_{j=i^*}^{i-1} \left(1+\|\rho_j^{(n)}\|\right) \le 2M_n \prod_{j=i^*}^{i-1} \left(1+2\int_{\frac{j}{2^n}}^{\frac{j+1}{2^n}} \|P(t)\|\,dt\right) \le$$

$$\le 2M_n \left(\frac{1}{i-i^*}\sum_{j=i^*}^{i-1}\left(1+2\int_{\frac{j}{2^n}}^{\frac{j+1}{2^n}} \|P(t)\|\,dt\right)\right)^{i-i^*} =$$

$$= 2M_n \left(1+\frac{2}{i-i^*}\int_{\frac{i^*}{2^n}}^{\frac{i}{2^n}} \|P(t)\|\,dt\right)^{i-i^*} \le$$

$$\le 2M_n \left(1+\frac{2}{i-i^*}\int_0^1 \|P(t)\|\,dt\right)^{i-i^*} < 2M_n e^{2\int_0^1 \|P(t)\|\,dt}$$

for all $i > i^*$. In the same way we also obtain this result for $i < i^*$, and for $i = i^*$ is true since $\|z_n(\frac{i^*}{2^n})\| \le 2M_n$. Therefore, we proved that

$$\|z_n(x)\| \le 2M_n e^{2\int_0^1 \|P(t)\|\,dt} \qquad (x \in [0,1[),$$

except for finitely many values of $n$. $M_n$ tends to zero as $n \to \infty$ which implies that $\|z_n(x)\|$, the second addend of the right hand side of (5.3), also tends uniformly to zero. Thus, the absolute difference $\|y(x) - \overline{y}_n(x)\|$ tends uniformly to zero. Consequently we obtain the statement of Teorem 2.2.

Finally, we prove the lemmas used in this section.

**Lemma 5.1.** *Let $i^*$ be the positive integer such that $\frac{i^*}{2^n} \le \xi < \frac{i^*+1}{2^n}$ holds, $i = 0, 1, \dots, 2^n - 1$ and $\rho_i^{(n)}$ the matrices defined in (5.6). Then, except for finitely many values of $n$ we have*

$$z_n\Big(\frac{i^*}{2^n}\Big) = \Big(\mathcal{I} + \frac{d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\Big)^{-1} m_n\Big(\frac{i^*}{2^n}\Big),$$

$$z_n\Big(\frac{i}{2^n}\Big) = \Big(\mathcal{I} + \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\Big)^{-1}\Big(m_n\Big(\frac{i}{2^n}\Big) - \sum_{j=i^*}^{i-1}\Big(\prod_{s=j+1}^{i-1}\Big(\mathcal{I} - \rho_s^{(n)}\Big)\Big)\rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big)\Big)$$

*for all $i > i^*$, and*

$$z_n\Big(\frac{i}{2^n}\Big) = \Big(\mathcal{I} - \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\Big)^{-1}\Big(m_n\Big(\frac{i}{2^n}\Big) - \sum_{j=i+1}^{i^*}\Big(\prod_{s=j+1}^{i-1}\Big(\mathcal{I} - \rho_s^{(n)}\Big)\Big)\rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big)\Big),$$

*for all $i < i^*$.*

**Proof.** All functions $S_{2^n}p_{k,l}$ and $z_n$ are constants on the dyadic intervals of length $2^{-n}$. Hence by Lemma 3.1 we obtain the following cases from (5.5):

$$z_n\Big(\frac{i^*}{2^n}\Big) = m_n\Big(\frac{i^*}{2^n}\Big) - \frac{d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)z_n\Big(\frac{i^*}{2^n}\Big),$$

if $i > i^*$ then

$$z_n\Big(\frac{i}{2^n}\Big) = m_n\Big(\frac{i}{2^n}\Big) - \frac{d_n(\xi)+\frac{1}{2}}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)z_n\Big(\frac{i^*}{2^n}\Big) - \\ - \frac{1}{2^n}\sum_{j=i^*+1}^{i-1}\mathcal{S}_{2^n}p\Big(\frac{j}{2^n}\Big)z_n\Big(\frac{j}{2^n}\Big) - \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)z_n\Big(\frac{i}{2^n}\Big)$$

and if $i < i^*$ then

$$z_n\Big(\frac{i}{2^n}\Big) = m_n(\frac{i}{2^n}) - \frac{d_n(\xi)-\frac{1}{2}}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)z_n\Big(\frac{i^*}{2^n}\Big) + \\ + \frac{1}{2^n}\sum_{j=i+1}^{i^*}\mathcal{S}_{2^n}p\Big(\frac{j}{2^n}\Big)z_n\Big(\frac{j}{2^n}\Big) + \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)z_n\Big(\frac{i}{2^n}\Big).$$

Consequently

$$\Big(\mathcal{I} + \frac{d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\Big)\, z_n\Big(\frac{i^*}{2^n}\Big) = m_n\Big(\frac{i^*}{2^n}\Big), \tag{5.7}$$

if $i > i^*$ then

$$\Big(\mathcal{I} + \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\Big)\, z_n\Big(\frac{i}{2^n}\Big) = m_n\Big(\frac{i}{2^n}\Big) - \frac{d_n(\xi)+\frac{1}{2}}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)z_n\Big(\frac{i^*}{2^n}\Big) - \\ - \frac{1}{2^n}\sum_{j=i^*+1}^{i-1}\mathcal{S}_{2^n}p\Big(\frac{j}{2^n}\Big)z_n\Big(\frac{j}{2^n}\Big), \tag{5.8}$$

and if $i < i^*$ then

$$\Big(\mathcal{I} - \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\Big)\, z_n\Big(\frac{i}{2^n}\Big) = m_n\Big(\frac{i}{2^n}\Big) + \frac{\frac{1}{2}-d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)z_n\Big(\frac{i^*}{2^n}\Big) + \\ + \frac{1}{2^n}\sum_{j=i+1}^{i^*-1}\mathcal{S}_{2^n}p\Big(\frac{j}{2^n}\Big)z_n\Big(\frac{j}{2^n}\Big). \tag{5.9}$$

Formulae above give us the values of $z_n$ from $m_n$ and $\mathcal{S}_{2^n}p$ by iteration, provided that the matrices

$$\mathcal{I} + \frac{d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big) \qquad \text{and} \qquad \mathcal{I} \pm \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)$$

are non-singular. Fortunately, we proved in Section 5.1 that it is not true only for finitely many values of $n$. We suppose than $n$ is large enough for this to happen. Then, we obtain the statement of the lemma for $i^*$ directly from (5.7). We prove the case $i > i^*$ by induction. First note that it is true for $i = i^* + 1$. Indeed by (5.8) and (5.7) we obtain

$$\Big(\mathcal{I} + \frac{1}{2^{n+1}}\mathcal{S}_{2^n}p\Big(\frac{i^*+1}{2^n}\Big)\Big)\, z_n\Big(\frac{i^*+1}{2^n}\Big) = \\ = m_n\Big(\frac{i^*+1}{2^n}\Big) - \frac{d_n(\xi)+\frac{1}{2}}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)z_n\Big(\frac{i^*}{2^n}\Big) = \\ = m_n\Big(\frac{i^*+1}{2^n}\Big) - \frac{d_n(\xi)+\frac{1}{2}}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\, \Big(\mathcal{I} + \frac{d_n(\xi)}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\Big)^{-1} m_n\Big(\frac{i^*}{2^n}\Big) = \\ = m_n\Big(\frac{i^*+1}{2^n}\Big) - \rho_{i^*}^{(n)} m_n\Big(\frac{i^*}{2^n}\Big).$$

Supposing that the statement of the lemma holds for some $i > i^*$ we have

$$\Big(\mathcal{I} + \frac{1}{2^{n+1}} \mathcal{S}_{2^n} p\Big(\frac{i+1}{2^n}\Big)\Big)\, z_n\Big(\frac{i+1}{2^n}\Big) =$$

$$= m_n\Big(\frac{i+1}{2^n}\Big) - \frac{d_n(\xi) + \frac{1}{2}}{2^n} \mathcal{S}_{2^n} p\Big(\frac{i^*}{2^n}\Big) z_n\Big(\frac{i^*}{2^n}\Big) - \frac{1}{2^n} \sum_{j=i^*+1}^{i} \mathcal{S}_{2^n} p\Big(\frac{j}{2^n}\Big) z_n\Big(\frac{j}{2^n}\Big) =$$

$$= m_n\Big(\frac{i+1}{2^n}\Big) - \frac{d_n(\xi) + \frac{1}{2}}{2^n} \mathcal{S}_{2^n} p\Big(\frac{i^*}{2^n}\Big) z_n\Big(\frac{i^*}{2^n}\Big) -$$

$$- \frac{1}{2^n} \sum_{j=i^*+1}^{i-1} \mathcal{S}_{2^n} p\Big(\frac{j}{2^n}\Big) z_n\Big(\frac{j}{2^n}\Big) - \frac{1}{2^n} \mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big) z_n\Big(\frac{i}{2^n}\Big) =$$

$$= m_n\Big(\frac{i+1}{2^n}\Big) + \Big(\mathcal{I} + \frac{1}{2^{n+1}} \mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big)\Big)\, z_n\Big(\frac{i}{2^n}\Big) - m_n\Big(\frac{i}{2^n}\Big) - \frac{1}{2^n} \mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big) z_n\Big(\frac{i}{2^n}\Big) =$$

$$= m_n\Big(\frac{i+1}{2^n}\Big) - \sum_{j=i^*}^{i-1} \Big( \prod_{s=j+1}^{i-1} \big(\mathcal{I} - \rho_s^{(n)}\big) \Big) \rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big) - \frac{1}{2^n} \mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big) \times$$

$$\times \Big(\mathcal{I} + \frac{1}{2^{n+1}} \mathcal{S}_{2^n} p\Big(\frac{i}{2^n}\Big)\Big)^{-1} \Big( m_n\Big(\frac{i}{2^n}\Big) - \sum_{j=i^*}^{i-1} \Big( \prod_{s=j+1}^{i-1} \big(\mathcal{I} - \rho_s^{(n)}\big) \Big) \rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big) \Big) =$$

$$= m_n\Big(\frac{i+1}{2^n}\Big) - \sum_{j=i^*}^{i-1} \Big( \prod_{s=j+1}^{i-1} \big(\mathcal{I} - \rho_s^{(n)}\big) \Big) \rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big) -$$

$$- \rho_i^{(n)} \Big( m_n\Big(\frac{i}{2^n}\Big) - \sum_{j=i^*}^{i-1} \Big( \prod_{s=j+1}^{i-1} \big(\mathcal{I} - \rho_s^{(n)}\big) \Big) \rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big) \Big) =$$

$$= m_n\Big(\frac{i+1}{2^n}\Big) - \rho_i^{(n)} m_n\Big(\frac{i}{2^n}\Big) - (\mathcal{I} - \rho_i^{(n)}) \sum_{j=i^*}^{i-1} \Big( \prod_{s=j+1}^{i-1} \big(\mathcal{I} - \rho_s^{(n)}\big) \Big) \rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big) =$$

$$= m_n\Big(\frac{i+1}{2^n}\Big) - \rho_i^{(n)} m_n\Big(\frac{i}{2^n}\Big) - \sum_{j=i^*}^{i-1} \Big( \prod_{s=j+1}^{i} \big(\mathcal{I} - \rho_s^{(n)}\big) \Big) \rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big) =$$

$$= m_n\Big(\frac{i+1}{2^n}\Big) - \sum_{j=i^*}^{i} \Big( \prod_{s=j+1}^{i} \big(\mathcal{I} - \rho_s^{(n)}\big) \Big) \rho_j^{(n)} m_n\Big(\frac{j}{2^n}\Big).$$

which means that the statement of the lemma also holds for $i+1$. Therefore, it is valid for all $i > i^*$. The case $i < i^*$ can be proved in a similar way. This completes the proof of the lemma. ■

**Lemma 5.2.** *For all $n \in \mathbf{N}$ we have*

$$\|m_n(x)\| \leq M_n \qquad (x \in [0,1[),$$

*where*

$$M_n := \omega_n^{(1)} q + \|y\|_\infty \omega_n^{(1)} P + 2mL_n \omega_n y.$$

**Proof.** The definitions of $m_n$ in (5.4) consists of three parts. For the first we have

$$\left\|S_{2^n}\left(\int\limits_{\xi}^{\cdot}(S_{2^n}q(t)-q(t))\;dt\right)(x)\right\| \le \left\|S_{2^n}\left(\int\limits_{\xi}^{\cdot}|S_{2^n}q(t)-q(t)|\,dt\right)(x)\right\| \le$$

$$\le \left\|\int\limits_0^1 |S_{2^n}q(t)-q(t)|\,dt\right\| \le \omega_n^{(1)}q$$

for all $x\in[0,1[$. Recall that by the absolute value of a matrix we mean the matrix formed by the absolute value if its entries, even for vectors.

For the estimation of the second part recall that all members of the exact solution $y$ are continuous on the interval $[0,1]$. This means that they are bounded functions, so $\|y\|_\infty<\infty$, and each attains its maximum at a point of the interval $[0,1]$. Obviously $\|y(x)\|\le\|y\|_\infty$ for all $x\in[0,1]$. Then

$$\left\|S_{2^n}\left(\int\limits_{\xi}^{\cdot}(\mathcal{S}_{2^n}p(t)-P(t))y(t)\,dt\right)(x)\right\| \le \left\|S_{2^n}\left(\int\limits_{\xi}^{\cdot}|\mathcal{S}_{2^n}p(t)-P(t))y(t)|\,dt\right)(x)\right\| \le$$

$$\le \left\|\int\limits_0^1 |\mathcal{S}_{2^n}p(t)-P(t))|\cdot|y(t)|\,dt\right\| \le \|y\|_\infty\omega_n^{(1)}P.$$

for all $x\in[0,1[$. Before the estimation of the last part of $m_n$ note that

$$\int\limits_{\frac{j}{2^n}}^{\frac{j+1}{2^n}} S_{2^n}y(t)-y(t)\,dt=0 \qquad (j=0,1,\dots,2^n-1)$$

since this property is valid for all integrable functions. Besides, $\mathcal{S}_{2^n}p$ is constant on the all dyadic intervals of length $2^{-n}$. For this reason

$$\int\limits_{\frac{j}{2^n}}^{\frac{j+1}{2^n}} \mathcal{S}_{2^n}p(t)(S_{2^n}y(t)-y(t))\,dt=0 \qquad (j=0,1,\dots,2^n-1).$$

The formula above is used in the estimation. Then we have two cases with respect to the used notation $\frac{i}{2^n}\le x<\frac{i+1}{2^n}$ and $\frac{i^*}{2^n}\le\xi<\frac{i^*+1}{2^n}$. If $i>i^*$ then

$$S_{2^n}\left(\int\limits_{\xi}^{\cdot}\mathcal{S}_{2^n}p(t)(y(t)-S_{2^n}y(t))\,dt\right)(x)=2^n\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\xi}^{\tau}\mathcal{S}_{2^n}p(t)(y(t)-S_{2^n}y(t))\,dt\,d\tau=$$

$$=2^n\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\xi}^{\frac{i^*+1}{2^n}}\mathcal{S}_{2^n}p(t)(y(t)-S_{2^n}y(t))\,dt\,d\tau+2^n\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\frac{i}{2^n}}^{\tau}\mathcal{S}_{2^n}p(t)(y(t)-S_{2^n}y(t))\,dt\,d\tau=$$

$$=2^n\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\xi}^{\frac{i^*+1}{2^n}}y(t)-S_{2^n}y(t)\,dt\,d\tau+2^n\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\frac{i}{2^n}}^{\tau}y(t)-S_{2^n}y(t)\,dt\,d\tau$$

and hence

$$\left\| S_{2^n}\left(\int\limits_{\xi}^{\cdot} \mathcal{S}_{2^n}p(t)(y(t)-S_{2^n}y(t))\,dt\right)(x)\right\| \le$$

$$\le 2^n\left\|\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\right\|\left\|\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\xi}^{\frac{i^*+1}{2^n}}|S_{2^n}y(t)-y(t)|\,dt\,d\tau\right\|+$$

$$+\,2^n\left\|\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\right\|\left\|\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\frac{i}{2^n}}^{\tau}|S_{2^n}y(t)-y(t)|\,dt\,d\tau\right\| \le$$

$$\le 2^n\left\|\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\right\|\left\|\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\frac{i^*}{2^n}}^{\frac{i^*+1}{2^n}}|S_{2^n}y(t)-y(t)|\,dt\,d\tau\right\|+$$

$$+\,2^n\left\|\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\right\|\left\|\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}|S_{2^n}y(t)-y(t)|\,dt\,d\tau\right\| \le$$

$$\le \left\|\frac{1}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i^*}{2^n}\Big)\right\|\omega_n y+\left\|\frac{1}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\right\|\omega_n y \le 2mL_n\omega_n y.$$

The calculation for $i<i^*$ is very similar. Moreover, for $i=i^*$

$$S_{2^n}\left(\int\limits_{\xi}^{\cdot} \mathcal{S}_{2^n}p(t)(y(t)-S_{2^n}y(t))\,dt\right)(x)=2^n\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\xi}^{\tau}\mathcal{S}_{2^n}p(t)(y(t)-S_{2^n}y(t))\,dt\,d\tau=$$

$$=2^n\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\xi}^{\tau}(y(t)-S_{2^n}y(t))\;dt\,d\tau$$

and hence

$$\left\| S_{2^n}\left(\int\limits_{\xi}^{\cdot} \mathcal{S}_{2^n}p(t)(y(t)-S_{2^n}y(t))\,dt\right)(x)\right\| \le$$

$$\le 2^n\left\|\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\right\|\left\|\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\xi}^{\tau}|S_{2^n}y(t)-y(t)|\,dt\,d\tau\right\| \le$$

$$\le 2^n\left\|\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\right\|\left\|\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}\int\limits_{\frac{i}{2^n}}^{\frac{i+1}{2^n}}|S_{2^n}y(t)-y(t)|\,dt\,d\tau\right\| \le \left\|\frac{1}{2^n}\mathcal{S}_{2^n}p\Big(\frac{i}{2^n}\Big)\right\|\omega_n y \le mL_n\omega_n y.$$

Summarizing the estimation of the three parts we obtain the statement of the lemma. ∎

**Lemma 5.3.** *For all* $i = 0, 1, \ldots, 2^n - 1$ *we have*

$$\left\| \left( \mathcal{I} + \frac{d_n(\xi)}{2^n} \mathcal{S}_{2^n} p\left(\frac{i^*}{2^n}\right)\right)^{-1} \right\| \le 2 \qquad \textit{and} \qquad \left\| \left( \mathcal{I} \pm \frac{1}{2^{n+1}} \mathcal{S}_{2^n} p\left(\frac{i}{2^n}\right)\right)^{-1} \right\| \le 2$$

*except for finitely many values of* $n$.

**Proof.** As we have previously said in Section 5.1, the matrices above are "nearly identity" if the value of $n$ is large enough. This is due to the fact that the entries of the matrix $\frac{1}{2^{n+1}} \mathcal{S}_{2^n} p(\frac{i}{2^n})$ tend uniformly to zero, according to

$$\max_{\substack{0 \le i < 2^n \\ 1 \le k, l \le m}} \left\{ \frac{1}{2^n} \left| S_{2^n} p_{k,l}\left(\frac{i}{2^n}\right)\right| \right\} \to 0.$$

This means that for all $i = 0, 1, \ldots, 2^n - 1$

$$\left\| \frac{d_n(\xi)}{2^n} \mathcal{S}_{2^n} p\left(\frac{i^*}{2^n}\right)\right\| \le \frac{1}{2} \qquad \text{and} \qquad \left\| \frac{\pm 1}{2^{n+1}} \mathcal{S}_{2^n} p\left(\frac{i}{2^n}\right)\right\| \le \frac{1}{2} \tag{5.10}$$

except for finitely many values of $n$. For simplicity denote by $S$ one of the matrices in (5.10). Form the equality $(\mathcal{I} + S)^{-1} = \mathcal{I} - S(\mathcal{I} + S)^{-1}$ and the basic properties of the norm we have

$$\|(\mathcal{I} + S)^{-1}\| \le 1 + \|S\| \|(\mathcal{I} + S)^{-1}\|.$$

Thus,

$$\|(\mathcal{I} + S)^{-1}\|(1 - \|S\|) \le 1.$$

Since $\|S\| \le \frac{1}{2}$ we obtain

$$\|(\mathcal{I} + S)^{-1}\| \le \frac{1}{1 - \|S\|} \le 2.$$

This completes the proof of the lemma. ∎

## 6. Conclusion

The method designed by Chen and Hsiao for solving systems of linear differential equations of first-order with constant coefficients may be extended for equations with not constant coefficients, and also for systems of linear differential equations with discontinuous righthand sides if all functions in the equations are integrable. The results obtained by György Gát and the author of this paper for one equation are generalized for problems with more than one equations. The proposed numerical solution is a Walsh polynomial, i.e. a piecewise constant function on intervals of length $\frac{1}{2^n}$. This numerical solution exists, except for finitely many values of $n$, and it converge uniformly to the exact solution.

A modification of this numerical solution is proposed which basically consists of keeping the value of it at the middle of these intervals, and connecting these points by linear interpolation. This modification increase significantly the accuracy of the approximation. In several cases the approximation by Walsh polynomials gives us a maximal difference which is reduced by almost a half when the value of $n$ increases by one. However, in these cases the modified solution reduces the maximal difference by almost a quarter.

## References

[1] **Blachman, N.M.,** Sinusoids versus Walsh functions, *Proceedings of the IEEE*, **62(3)** (1974), 346–354.

[2] **Chen, C.F. and C.H. Hsiao,** A state-space approach to Walsh series solution of linear systems, *International Journal of Systems Science*, **6(9)** (1975), 833–858.

[3] **Chen, C.F. and C.H. Hsiao,** A Walsh series direct method for solving variational problems, *Journal of the Franklin Institute*, **300(4)**, (1975), 265–280.

[4] **Chen, C.F. and C.H. Hsiao,** Design of piecewise constant gains for optimal control via Walsh functions, *IEEE Transactions on Automatic Control*, **20(5)** (1975), 596–603.

[5] **Chen, C.F. and C.H. Hsiao,** Time-domain synthesis via Walsh functions, in *Proceedings of the Institution of Electrical Engineers*, **122** (1975), 565–570.

[6] **Chen, C.F. and C.H. Hsiao,** Walsh series analysis in optimal control, *International Journal of Control*, **21(6)** (1975), 881–897.

[7] **Corrington, M.,** Solution of differential and integral equations with Walsh functions, *IEEE Transactions on Circuit Theory*, **20(5)** (1973), 470–476.

[8] **Filippov, A.F.,** *Differential Equations with Discontinuous Righthand Sides*, Kluwer Academic Publishers Group, 1988.

[9] **Fine, N.J.,** On the Walsh functions, *Trans. Am. Math. Soc.*, **65** (1949), 372–414.

[10] **Gát, G. and R. Toledo,** A numerical method for solving linear differential equations via Walsh functions, in *Proceedings of the 18th International Conference on Computers (part of CSCC 2014), Santorini Island, Greece, July 17–21, 2014*, Advances in Information Science and Applications volume 2, 2014, 334–339.

[11] **Gát, G. and R. Toledo,** Estimating the error of the numerical solution of linear differential equations with constant coefficients via Walsh polynomials, *Acta Math. Acad. Paedagog. Nyházi.* (*N.S.*), **31(2)** (2015), 309–330.

[12] **Gát, G. and R. Toledo,** Numerical solution of linear differential equations by Walsh polynomials approach, *Studia Scientiarum Mathematicarum Hungarica*, **57(2)** (2020), 217–254.

[13] **Gibbs, J.E. and H.A. Gebbie,** Application of Walsh functions to transform spectroscopy, *Nature*, **224** (1969), 1012–1013.

[14] **Harmuth, H.F.,** *Transmission of Information by Orthogonal Functions. 2nd ed.*, Berlin-Heidelberg-New York: Springer-Verlag. XII, 393 p., 1972.

[15] **Nagy, K.,** The existence and unicity of numerical solution of initial value problems by Walsh polynomials approach, 2020
`https://arxiv.org/pdf/2006.14549.pdf`

[16] **Ohta, T.,** Expansion of Walsh functions in terms of shifted Rademacher functions and its applications to the signal processing and the radiation of electromagnetic Walsh waves, *IEEE Transactions on Electromagnetic Compatibility*, **EMC-18** (1976), 201–205.

[17] **Rao, G.P.,** *Piecewise Constant Orthogonal Functions and their Application to Systems and Control*, volume 55, Springer, Cham, 1983.

[18] **Schipp, F., W.R. Wade and P. Simon,** *Walsh Series. An Introduction to Dyadic Harmonic Analysis*, Adam Hilger, Bristol and New York, 1990.

[19] **Shih, Yen-Ping and Joung-Yi Han,** Double walsh series solution of first-order partial differential equations. *International Journal of Systems Science*, **9(5)** (1978), 569–578.

**R. Toledo**
Department of Numerical Analysis
Faculty of Informatics
Eötvös Loránd University
H-1117 Budapest, Pázmány Péter sétány 1/C
Hungary
`toledo@inf.elte.hu`